

Reinhold Schmitt

DIE GESPRÄCHSPAUSE: Verbale „Auszeiten“ aus multimodaler Perspektive

Abstract

Der folgende Beitrag hat zwei Schwerpunkte: Zum einen wird die interaktive Struktur einer 12 Sekunden dauernden Pause und deren soziale Bedeutung in einem spezifischen Kontext aus gesprächsanalytisch-gesprächsrhetorischer Perspektive analysiert. Damit wird ein interaktives Phänomen zum Gegenstand der Untersuchung, das – aus triftigen Gründen und in motivierter Weise – in der bisherigen gesprächsanalytischen Forschung einen eher peripheren Stellenwert besaß: der temporäre Verzicht der Interaktionsbeteiligten auf verbalen Ausdruck.

Zum anderen wird diese Abwesenheit von Verbalität zum Anlass genommen, einen Aspekt zu thematisieren, der sich meines Erachtens auf die künftige Entwicklung der Gesprächsanalyse gravierend auswirken wird: die Tatsache, dass die empirische Grundlage gesprächsanalytischer Untersuchungen immer häufiger aus audiovisuellen Daten besteht. Diese „visuelle Revolution“, die durch die technologische Entwicklung in Form problemlos einsetzbarer Videokameras angestoßen wird, hat weit reichende Konsequenzen für die Theorie, Methodologie und Methode eines Analyseansatzes, der sich bislang und programmatisch bei der Analyse interaktiver Ordnungsstrukturen primär auf Verbalität konzentriert hat.

The following article has two main focuses: Firstly, the interactive structure of a 12 second continuous pause and its social meaning in a specific context is analysed from the perspective of conversation analysis and conversational rhetoric. The investigation thus focuses on an interactive phenomenon, which - for good reasons and in a motivated fashion - has had a rather peripheral value in past research into conversation analysis: the temporary renunciation of verbal expression on the part of the participants in the interaction.

Secondly, this absence of verbal expression is used to focus on an aspect which in my judgement will seriously affect the future development of conversation analysis: the fact that the empirical basis of investigations in this field consists to an ever increasing degree of audiovisual data. This "visual revolution", which has been encouraged by technological progress in form of easily usable video cameras, has far-reaching consequences for the theory, methodology and method of an analytical approach which so far has concentrated primarily and programmatically on verbal exchanges in the analysis of interactive structures.

1. Pausen als Interaktionsphänomene

Bei der Besprechung einer Filmidee, die eine vierköpfige Gruppe von Filmstudenten zusammen mit ihrem Lehrer durchführt, kommt es – nachdem die Studenten ihren Plan für die Ausgestaltung einer zentralen Szene detailliert ausgeführt haben – zu einer kritischen Nachfrage des Lehrers. Er reformuliert zunächst den studentischen Vorschlag und problematisiert anschließend die mit dem Vorschlag verbundene Mehrdeutigkeit der Szene. Zum Abschluss seiner Problematisierung fordert er die Studenten in thematisch unspezifischer, jedoch erkennbar relevanzhochgestufter Weise zu einer Antwort auf. Als Reaktion auf seine Aufforderung entsteht eine zwölf Sekunden dauernde Pause¹ (*12*). Diese Pause wird als zentrales Interaktionsphänomen im Fokus der nachfolgenden Ausführungen stehen.

¹ Die verwendete Kurzform „Pause“ ist im Folgenden immer als *Sprech-* und nicht als *Interaktionspause* zu verstehen.

Pausen, wie die oben beschriebene, die systematisch an *transition relevance places* zu beobachten sind, haben als Untersuchungsgegenstand in der Gesprächsforschung keine bedeutende Rolle gespielt – sieht man von vereinzelt Arbeiten ab (Umiker-Sebeok (1980), Bergmann (1982), Jefferson (1989), Müller/Paul (1999)). Im Bereich der Psycholinguistik waren sie von Interesse als Teil der Untersuchung von Verzögerungsphänomenen bei der Erforschung der Sprachproduktion (Rochester (1973), Butterworth (1975), Seyfeddinipur/Kita (2001)).

Die Beschreibung und Konzeptualisierung solcher turnvakanten Pausen² als Untersuchungsgegenstand der Gesprächsanalyse sind abhängig von der Perspektive, aus der sie erfolgen. Die Perspektive wiederum ist teilweise historisch indexikal: Mehr als bei verbalen Phänomenen zeigt sich bei der Analyse von turnvakanten Pausen, dass die Gegenstandskonstitution nicht nur eine Frage der Theorie und des Untersuchungsinteresses ist. Sie ist ebenso technisch bedingt, d.h. sie wird durch die zur Verfügung stehenden Dokumentationsmedien geprägt.

In den Anfängen der Gesprächsanalyse wurden für Untersuchungen primär verbale Daten in Form von Tonaufnahmen und deren Verschriftlichung herangezogen. Die technischen Bedingungen führten dabei zu einer – zwar nicht theoretischen, aber doch analysefaktischen – Priorisierung des Verbalen gegenüber anderen körperlichen Ausdrucksformen: Gesprächsanalyse beschäftigte sich mit der Analyse des Gesprochenen.³

Bergmann (1982) hat exemplarisch vorgeführt, was man bei einer ausschließlich auf verbalen Daten gegründeten Untersuchung über turnvakante Pausen methodisch kontrolliert aussagen kann. Er verdeutlichte, dass bei der Rekonstruktion von Pausen (er selbst spricht von *Schweigephasen*) die jeweils konkrete Platzierung im Interaktionsverlauf die zentrale Erklärungsressource darstellt. Die interaktive Bedeutung einer Pause ergibt sich demnach primär auf Grund ihres vorgängigen Kontextes und des nachfolgenden interaktiven Geschehens. Die Pause hat dabei den konstitutionslogischen Status eines Paarteils und kann als „second pair part“ bestimmt werden. Man kann dann auf der Basis des Sequenzierungsmechanismus analytisch entscheiden, wer die Pause „verantwortet“ und warum sie auf die vorliegende Weise beendet wird.

Kommt es in einem mit Video dokumentierten Gespräch zu einer längeren Schweigephase, verändert sich auf Grund der Medialität die konstitutive Qualität des Phänomens und dessen kategorialer Zuschnitt: Es handelt sich nun um eine Phase, in der die Beteiligten zwar verbal abstinent sind, dafür aber gestisch-körperlich ausgesprochen aktiv sein können.

Von der „verbalen“ Pausenperspektive bleiben jedoch der paardefinierte Status des Phänomens und das damit verbundene methodische Vorgehen seiner sequenziellen Rekonstruktion erhalten. Neben der detaillierten Rekonstruktion des vor- und nachgängigen

² „Turnvakante Pausen“ im Sinne von Bergmann (1982) entstehen nicht bei und als Teil der Produktion von Äußerungen, sondern tauchen an Stellen möglichen Sprecherwechsels auf.

³ Die explizite Formulierung der holistischen Qualität von Interaktion erfolgt erst spät: „Rather than looking at the talk, intonation, and body movement [...] as different channels of behavior to be analyzed separately, it seems more profitable to conceptualize what is happening as a single, interactive activity of assessment that the participants collaboratively recognize, bring to a climax or peak, and then withdraw from.” (Goodwin 1996, S. 371); siehe auch Dausendschön-Gay/Krafft (2002).

Geschehens kommt nun aber als zentrale Aufgabe die sequenzielle Rekonstruktion des sichtbaren Geschehens in der Phase verbaler Abstinenz selbst hinzu: *die interaktive und sequenzielle Struktur der Pause*.

Durch das Medium bedingt ergibt sich also nunmehr die Möglichkeit, Pausen als relevante interaktive Phänomene nicht nur – wie bisher in Transkripten – festzuhalten, sondern hinsichtlich ihrer internen Struktur zu beschreiben. Es eröffnet sich eine analytische Binnenperspektive⁴, und die zentrale Aufgabe besteht nun darin, die gestisch-körperliche Bearbeitung einer durch den vorgängigen Kontext etablierten interaktiven Aufgabe zu rekonstruieren (in meinem Fall ist das die durch eine Frage etablierte Reaktionsverpflichtung). Die Aktivitäten der Beteiligten während der verbalen Auszeit werden also als konstitutiver Bestandteil und erster Teil der Bearbeitung eines aktuellen interaktiven Problems begriffen. Folglich geht es in der Analyse darum, die sichtbaren Verfahren und Techniken zu rekonstruieren, welche die Beteiligten unter den spezifischen Kontextbedingungen einsetzen, um ein aktuelles interaktives Problem unter Verzicht auf verbalen Ausdruck zu lösen.

Betrachtet man die Forschungen zum Verhältnis von verbalem und körperlichem Verhalten in Kontexten, die für die hiesige Auseinandersetzung relevant sind, ergibt sich folgendes Bild: Zunächst einmal sind Beiträge aus dem Forschungsfeld der *Kinesik* zu nennen, die sich auf die Beschreibung und Konzeptualisierung der eigenständigen strukturellen, funktionalen und interaktiven Eigenschaften körperlichen Verhaltens in Anlehnung an eine sprachwissenschaftliche Perspektive konzentrieren (vergleiche beispielsweise Birdwhistell (1952, 1970 und 1984)). „Kinesics is the study of the visual aspects of non-verbal, interpersonal communication“ (Birdwhistell 1960, S. 54).

Weitere wichtige Bezugspunkte sind Arbeiten, die das Verhältnis von verbalem und körperlichem Verhalten von einem methodologischen Standpunkt aus untersuchen, der als *context analysis* bezeichnet wird (zum Beispiel Kendon (1972, 1980 und 1994) und Scheflen (1963, 1964, 1972 und 1973), der das Verhältnis stärker aus einer Perspektive kultureller Determination betrachtet). Dieser Ansatz wird charakterisiert durch die „emphasis it places on the importance of examining the behavior of people in interaction in the context they occur“ (Kendon 1990, S. 15).

Inzwischen gibt es eine beachtliche Anzahl von Untersuchungen zum Zusammenhang speziell von Rede und Gestik, in denen analysiert wird, wie Gesten mit Äußerungen verbunden sind, um bestimmte Aspekte des Gesagten zu verstärken, hervorzuheben und zu symbolisieren (ein guter Überblick hierzu ist Müller (1998)). Ein weiterer Schwerpunkt in diesem Forschungsfeld betrifft die Beziehung von Sprache und Kognition (beispielsweise McNeill (1992 und 2000)).

Auch im Bereich der Konversationsanalyse gibt es einige Arbeiten, die das Verhältnis von verbalem und körperlichem Verhalten zum Gegenstand haben (Schegloff (1984); Heath (1992), Müller (1998)). Ebenso liegen Analysen aus konversationsanalytischer Perspektive vor, die sich mit den unterschiedlichen Aufgaben beschäftigen, die Gesten, Blickorganisation und Körperpositur in der Interaktion übernehmen: beispielsweise für

⁴ Diese kann bei einer rein verbalen Datengrundlage nicht eingenommen werden, da in der Aufnahme nichts zu hören und im Transkript außer der Pausenmarkierung nichts zu lesen ist.

die Signalisierung von Beteiligungsweise (Goodwin (1981), Heath (1982, 1986)), die Organisation von Wortsuche (Goodwin/Goodwin (1986)), die Herstellung gemeinsamer Orientierung (Goodwin (1986)), Korrektur und Unterbrechungen (Fornel (1991)), die Turn-Organisation (Streeck/Hartge (1992)), die Einladung von Höreraktivitäten (Heath (1992)), die Projektion von Handlungen (Streeck (1995)), die Organisation von Hörerverhalten (Streeck (1994)), die Konstitution narrativer Strukturen (Müller (2003)) oder die Markierung von Beteiligungsweisen und der Segmentierung handlungsrelevanter Einheiten (Dausendschön-Gay/Krafft (2002)).

Es lassen sich jedoch kaum konversationsanalytische Untersuchungen finden, die analysieren, wie – und warum – Gesprächsbeteiligte in bestimmten Phasen bei der Bearbeitung interaktiver Aufgaben gänzlich auf das verbale Medium verzichten und für eine gewisse Zeit ausschließlich gestisch-körperlich tätig werden.⁵ Der vorliegende Beitrag beschäftigt sich genau mit diesem Fall: Es geht darum, wie Interaktionsbeteiligte (hier: Filmstudenten) bei der Bearbeitung eines lokalen interaktiven Problems (hier: Wer soll auf die Frage des Dozenten antworten?) *kinesische Turns*⁶ einsetzen und nicht auf *verbale Äußerungen* zurückzugreifen.

Die Vorstellung von kinesischen Turns ist keine neue Idee. Sie taucht bereits – wenn auch nicht unter diesem Namen – in einem der klassischen konversationsanalytischen Aufsätze auf. So findet man beispielsweise bei Schegloff (1972, S. 358) folgende Stelle:

“It is to be noted that a summons occurs at the first part of a two part sequence. Just as there are various items that can be used as summonses, so are there various items that are appropriately used as answers, e.g., “Yes?”, “What?”, “Uh huh?”, *turning of the eyes or of the body to face the beckoner*, etc. Some typical summons-answer sequences are: telephone ring – “hello”; “Johnny?” – “yes?”; “Excuse me” – “Yes”; “Bill?” – *looks up*.” (Hervorhebungen von R.S.)

Diese Passage ist insofern interessant, weil sie die Konstitution von Aufforderungs-Antwort-Sequenzen (so genannte *adjacency pairs*) zeigt, die aus zwei konditionell relevant aufeinander bezogenen Teilen bestehen: Einer dieser Teile ist verbal produziert, der andere durch Körperverhalten (d.h. es handelt sich bei letzterem um einen kinesischen Turn). Es werden zwei Beispiele angeführt, die sich im Hinblick auf die Platzierung des kinesischen Turns unterscheiden. Das erste zeigt den kinesischen Turn (die Antwort) in der reaktiven Position „*turning of the eyes or of the body to face the beckoner*“ und „*looks up*“ als Reaktion auf die Aufforderung „Bill?“

Das zweite Beispiel zeigt den kinesischen Turn (die Aufforderung) in initiativer Position. Strenggenommen ist es nicht wirklich die Aufforderung (dies wäre das Wählen der Telefonnummer), sondern ihre akustischen Konsequenzen, nämlich das Klingeln des

⁵ Siehe jedoch Müller/Paul (1999).

⁶ Ich benutze den Begriff „kinesischer Turn“ in bewusster Anlehnung an den in der Gesprächsforschung etablierten klassisch-verbale Turnbegriff. Ich bezeichne damit pragmatisch-sequenziell beschreibbare körperliche Verhaltensweisen (Gestikulation, Mimik, Veränderung der Körperpostur und Blickveränderung), die unter bestimmten interaktionsstrukturellen Bedingungen verbale Äußerungen äquivalent ersetzen können bzw. diesen gegenüber gesprächsrhetorisch im Vorteil sind. Der Begriff hat den Status eines Arbeitsbegriffs. Als solcher trägt er einerseits der Tatsache Rechnung, dass Interaktionsbeteiligte nicht nur mit verbalen Turns interaktiv handeln und zur Etablierung der interaktiven Ordnung beitragen. Andererseits ist er ein Fingerzeig, der auf die Frage nach dem Verhältnis zum klassisch-verbale Turn in theoretischer und methodischer Hinsicht als anstehende Aufgabe hinweist. Weitere Ausführungen zum Konzept „kinesischer Turn“ finden sich in Schmitt (i.V. b).

Telefons. Relevant ist hier nur eines: Der initiative Zug wird nicht verbal vollzogen. Es wird also deutlich, dass bereits in den frühen Tagen der Konversationsanalyse die Produktion von Turns nicht auf Verbalität beschränkt war, sondern auch Platz ließ für kinesische Varianten. Bei der weiteren Entwicklung der Konversationsanalyse und ihrer Fokussierung auf *talk-in-interaction* als ihrem zentralen Gegenstand, spielten kinesische Turn-Varianten keine wesentliche Rolle mehr.

Sprechpausen sind keine interaktiven Auszeiten.⁷ Nicht nur, dass sich die Interaktionsbeteiligten auch in Phasen verbaler Abstinenz im Verständnis von Goffman (1961) weiterhin in wechselseitiger Wahrnehmung und fokussierter Interaktion befinden. Die Beteiligten kommunizieren – wie sich noch zeigen wird – auch in den Pausen weiter, d.h. auch die formale *turn-by-turn*-Entwicklung und deren sequenzielle Organisation unter Benutzung der *turn-taking*-Maschine geht weiter.⁸ Die Maschine läuft zwar lautlos, aber durchaus nicht weniger strukturiert, sondern systematisch, effektiv und interaktiv folgenreich.

An dem Beispiel, das ich analysieren werde, sieht man, dass es sogar recht betriebsam in dieser zwölfsekündigen Interaktionsphase zugeht: Es findet nämlich eine komplexe Aushandlung zwischen den Beteiligten statt. Nicht nur, dass die Beteiligten hierzu in spezifischer Weise beitragen, man kann sogar die einzelnen Beiträge hinsichtlich ihrer pragmatischen und sequenziellen Spezifik und den damit jeweils verbundenen Anschlussimplikationen präzise rekonstruieren.

Auch wenn die zwölfsekündige Pause unter sehr spezifischen Kontextbedingungen zu Stande kommt, möchte ich – von ihrer Analyse ausgehend – dafür argumentieren, die für die Gesprächsanalyse noch immer charakteristische Konzentration auf Verbalität aufzugeben und alle für die Interaktion relevanten Ausdrucks- und Handlungsebenen systematisch zu berücksichtigen.⁹ Die Suche nach einem angemessenen Begriff für eine solche systematische Erweiterung gestaltet sich jedoch nicht einfach.

Auch der Begriff „holistisch“ macht diesbezüglich keine Ausnahme: Er suggeriert, man könne mit der Erweiterung der Verbalität die Totalität dessen, was wir als Beteiligte in der Kommunikation alles wahrnehmen, tatsächlich dokumentieren. In dieser Hinsicht bleibt der holistische Anspruch jedoch eine relative Größe. Bei der Dokumentation einer Situation mit Video werden immer noch viele Informationen nicht erfasst, abgeschnitten

⁷ „Die kommunikative Interaktion hört nicht auf, wenn die Partner in Schweigen versinken, und fängt nicht neu an, wenn sie wieder Laute erzeugen; andere Kommunikationsspuren bleiben auch dann in Betrieb, wenn die Gehörspur nicht benutzt wird“ Birdwhistell (1984, S. 193).

⁸ Die Maschinen-Metapher wird in klassischen Arbeiten der Konversationsanalyse benutzt, um den formalen Mechanismus zu benennen, den die Interaktionsbeteiligten – unabhängig von ihren individuellen Zielen – benutzen müssen, um geordnete Kommunikation zu etablieren und aufrecht zu erhalten. „... two basic features of conversation are proposed: (1) at least, and no more than, one party speaks at a time in a single conversation; and (2) speaker change recurs. The achievement of these features singly, and especially the achievement of their co-occurrence, is accomplished by conversationalists through the use of a ‘**machinery**’ (Hervorhebung von R.S.) for ordering speaker turns sequentially in conversation. The turn-taking machinery includes as one component a set of procedures for organizing the selection of ‘next speaker’, and as another, a set of procedures for locating the occasions on which transition to any next speaker may or should occur“; Schegloff/Sacks (1973, S. 293).

⁹ Ich schließe mich hier konzeptionellen Überlegungen von Dausendschön-Gay/Krafft (2002) an.

oder beispielsweise durch die Kameraführung ausgeblendet bzw. vor-interpretiert.¹⁰ Eine holistische Vorstellung von Kommunikation ist also aus der Sicht ihrer Dokumentation immer nur eine Annäherung und letztlich weitgehend durch die technischen Bedingungen vorgegeben und beschränkt.

Als Alternative hat sich inzwischen der Begriff „multimodale Kommunikation“ etabliert.¹¹ Unter multimodaler Kommunikation wird eine Konzeption verstanden, die davon ausgeht, dass es für Interaktionsbeteiligte unterschiedliche Modalitäten gibt, sich in kommunikationsrelevanter Weise auszudrücken, Handlungsziele zu erreichen, soziale Bedeutung zu konstituieren und alle möglichen Arten interaktiver Arbeit zu betreiben. Hierzu zählen beispielsweise: Verbalität, Prosodie, Mimik, Gestik, Körperpositur, Körperkonstellation und Blickverhalten.

Die genaue Anzahl interaktionsrelevanter Modalitäten spielt für den hiesigen Darstellungszusammenhang zunächst keine Rolle. Es geht vielmehr aus einer methodisch-methodologischen Perspektive heraus zunächst darum, die einzelnen Modalitätsebenen – neben ihrer wechselseitigen Vernetztheit – gerade auch in ihrer empirischen Eigenständigkeit zu begreifen. Eine solche Fokussierung einzelner Modalitätsebenen und der für sie charakteristischen Verhaltens- und Ausdrucksaspekte ermöglicht zweierlei: Zum einen rückt die spezifische Leistung einzelner Modalitätsebenen für die Strukturierung von Kommunikation in den Mittelpunkt des Erkenntnisinteresses. Zum anderen wird die Notwendigkeit deutlich, für die methodisch kontrollierte Beschreibung einzelner Modalitätsebenen angemessene Verfahren und Konzepte zu entwickeln.

Der letzte Punkt impliziert, dass man die für die Untersuchung verbalen Ausdrucks entwickelten Analyseverfahren und Konzepte nicht fraglos auf die anderen Modalitätsebenen übertragen kann. Vielmehr muss die Frage ihrer Übertragbarkeit und Adäquatheit methodisch reflektiert und empirisch beantwortet werden.

2. Die Situation „Pitching in Teams“

Das Beispiel, um das es im Folgenden gehen wird, stammt aus dem „Hamburger Filmkorpus“. Diese umfangreiche Sammlung von Videoaufzeichnungen beinhaltet zum einen Aufnahmen, welche die Arbeit an unterschiedlichen Drehorten dreier Filmteams zeigen (so genannte Set-Aufnahmen).¹² Zum anderen besteht sie aus Arbeitssitzungen (so genannte Pitching-Sitzungen), in denen Gruppen von Filmstudenten gemeinsam mit ihren Dozenten Filmideen entwickeln.¹³

Bei der Pitching-Sitzung, aus der das Beispiel stammt, treffen die Studenten zum ersten Mal als Gruppe mit den Dozenten zusammen, um Ideen für ihre ersten Filme zu präsen-

¹⁰ Zu dieser Problematik siehe die methodologischen Überlegungen in Schmitt (2001).

¹¹ Siehe beispielsweise die Mannheimer Treffen zu „Fragen der multimodalen Kommunikation“; Bericht in Schmitt (2004).

¹² Diese Set-Aufnahmen haben Daniela Heidtmann und ich im Rahmen unserer Kooperation mit dem Filmstudium der Universität Hamburg gemacht. Die Kooperation mit dem Filmstudium Hamburg ist – aus anwendungsbezogener Perspektive – dokumentiert in Schmitt/Heidtmann (2003 a, b); unter einem primär wissenschaftlichen Erkenntnisinteresse liegen erste Ergebnisse vor in Schmitt (i.V.a).

¹³ Dieser Teil wird von Daniela Heidtmann im Rahmen ihrer Dissertation „Kooperative Prozesse im Lehr-Lern-Diskurs“ ausgewertet.

tieren. Die Gruppenmitglieder kennen sich zu diesem Zeitpunkt erst seit zwei Monaten und konnten sich ihre Mitstreiter/innen nicht aussuchen, sondern wurden „zwangsverheiratet“. Die Konstellation ist jedoch in einer Hinsicht für alle Gruppen gleich: Sie sind jeweils zu viert und setzen sich aus je einem Vertreter der Klassen „Drehbuch“, „Regie“, „Produktion“ und „Kamera“ zusammen.

Die zentrale Aufgabe der Gruppe besteht darin, ihre Ideen zu „pitchen“, d.h. in wenigen Sätzen auf den Punkt zu bringen, um was es in ihrer Geschichte geht: Wer ist der Protagonist, welches Ziel hat er, welches Hindernis muss er bei der Erreichung seines Zieles überwinden und was sind die Kosten für ihn, wenn er an dem Hindernis scheitert?

Bei genauer Betrachtung stehen für die Studenten noch ganz andere Dinge auf dem Spiel: die gemeinsame Filmidee, die Präsentation der Gruppe als kreatives, funktionierendes Team und (in latenter Konkurrenz zu den Anderen) die Selbstdarstellung des Einzelnen als Individuum mit künstlerischem Anspruch. Die Situation hat zudem den Charakter einer Prüfung, bei der die eingereichte Geschichte als kollektive Leistung bewertet wird.¹⁴

Insgesamt stellen sich den Studenten sehr unterschiedliche, eng miteinander verwobene, teilweise widersprüchliche Anforderungen. Diese müssen sie in einem manifest hierarchisch strukturierten Rahmen bearbeiten, für den latente Fremdheit ein wesentliches Merkmal der Beziehungsstrukturen sowohl innerhalb der Gruppe als auch gegenüber den Dozenten darstellt.

3. Der relevante Kontext

Zu diesen situativen Bedingungen gesellen sich interaktionsstrukturelle Aspekte, die den Druck auf die Studenten weiter verstärken. Dies wird deutlich, wenn man sich den für das Zustandekommen und die interaktive Bedeutung der Sprechpause relevanten Kontext vergegenwärtigt.

Die Gruppe will einen Film über einen Medizinstudenten (Alex) drehen, der in seinem Studium zu scheitern droht, weil er keine Leichen sezieren kann. Die Stoffentwicklung ist an der zentralen Szene angelangt, in der Alex mit einem Skalpell über der Leiche steht und nicht in der Lage ist, diese zu öffnen. Um sein spezifisches Problem zu verdeutlichen, will die Gruppe den unproblematischen Umgang der anderen Medizinstudenten mit den Leichen als Fließbandtätigkeit darstellen: Bei einer Kamerafahrt soll gezeigt werden, dass diese quasi „industriell“ bearbeitet werden.

Ralf (RA), der Dozent, macht mit seiner Äußerung *das hab ich jetzt noch nicht-* deutlich, dass er mit den vorangegangenen Ausführungen zur industriellen Variante Verstehensprobleme hat. In Reaktion hierauf skizziert zunächst Anja diese Idee erneut, wobei sie punktuell von Silvia unterstützt wird. Im weiteren Verlauf übernimmt Özkan (ÖZ) von Anja die Darstellungsinitiative und konzentriert sich auf die konkrete Ausgestaltung der Szene „Alex mit dem Skalpell vor der Leiche“. Er beendet seinen Beitrag damit, dass er den dramaturgischen Effekt seiner Darstellung hervorhebt:

¹⁴ Einer der Studenten bezeichnet die Situation als „Tribunal“. Die wesentlichen Strukturelemente der Pitching-Situation und ihre interaktiven Auswirkungen sind detailliert beschrieben in Schmitt/Heidtmann (2002 b).

ÖZ: dramaturgischer aspekt wäre noch (... ..)

ÖZ: herzklopfen oder so noch zu hören- ** und dass er das

ÖZ: gefühl hat * vielleicht klopft es aus der leiche

ÖZ: raus oder so keine ahnung dass |man das irgendwie |
RA: |na gut aber wenn ihr|

Überlappend mit dem Abschluss der Äußerung von Özkan schaltet sich der Dozent wieder ein. Sein Beitrag besteht aus zwei größeren, pragmatisch eigenständigen, jedoch aufeinander bezogenen Teilen. Im ersten Teil fasst er zunächst noch einmal die von den Studenten entworfene Szene zusammen. Im Anschluss daran fragt er nach den Implikationen und der Bedeutungsvielfalt des entworfenen Bildes „Alex mit dem Skalpell vor der Leiche“. Seine Frage thematisiert das Spektrum dessen, was alles in diesem Bild steckt: *was kann das den alles/ was erzählt denn das alles ** >dieses bild< *I**.

Der zweite Teil seines Beitrags ist als Verdeutlichung des ersten markiert (*also*). Dies erweckt den Eindruck, für die Beantwortung seiner Frage durch die Studenten seien noch Detaillierungen seinerseits notwendig.

RA: |also we"nn er-| * zi"ttert herzklopfen *
K |"STICHT" MIT ZEIGEFINGER IN RICHTUNG ÖZ|

RA: schwei"ß und so wei"ter ** |dann |weiß| ich
K | MIT
K |SCHULTERZUCKEN|

RA: doch noch nicht ob es * ob er irgendwie in einem
K NACHDRUCK GESPROCHEN

RA: verhä"ltnis zu dieser person steht| * die da
K |

RA: gestorben ist * zu dieser leiche steht oder↑ *12*

Bei genauem Hinsehen handelt es sich jedoch weniger um eine Verdeutlichung als um eine Neufokussierung, die seiner Frage nun eine sehr konkrete Richtung gibt. Sie zielt nicht mehr in die Breite, sondern wird punktuell auf *eine* der möglichen Bedeutungsimplicationen der Szene zugespielt: Das entworfene Bild könne den Eindruck vermitteln, die Schwierigkeiten, die Alex mit dem Öffnen der Leiche habe, hingen damit zusammen, dass er eine Beziehung zu der Leiche habe (er die Person also zu Lebzeiten kannte).

Die Äußerung *dann weiß ich doch noch nicht ob es * ob er irgendwie in einem verhä"ltnis zu dieser person steht* zeigt folgenden interessanten Punkt: Die Formulierung „noch nicht wissen (ob)“ impliziert, dass noch Voraussetzungen bzw. Informationen fehlen, damit aus der Wissensunsicherheit Gewissheit werden kann. Die Äußerung des Dozenten kann also so verstanden werden, als müsse (oder solle) Alex tatsächlich eine Beziehung zu der Leiche haben. Für die Studenten wäre damit die Notwendigkeit verbunden, genau diesen Punkt szenisch zu verdeutlichen. Der Dozent beendet seinen Bei-

trag mit der als *tag question* realisierten Aufforderung *oder* ↑ an die Studenten und markiert damit deutlich einen Übergabepunkt.

Der Dozent engagiert sich in dem zitierten Ausschnitt auf der Grundlage einer Problemmatisierungsorientierung: Er weist die Studenten sofort auf relevante und kritische Implikationen ihrer Entwürfe und Ideen hin. Hierdurch wird jedoch seine grundsätzliche Würdigung der Ausführungen von Özkan überdeckt. Diese Würdigung findet ihren Ausdruck durch mehrfaches, manifestes Kopfnicken, hinterlässt verbal jedoch keine Spuren. Die Äußerung des Dozenten schafft für die Gruppe Anschlussbedingungen, die durch eine starke konditionelle Relevanz gekennzeichnet sind bei gleichzeitiger Unsicherheit, was inhaltlich der relevante Anforderungskern ist.

Die Studenten reagieren hierauf mit einem längeren Schweigen und einer gestisch-körperlichen Aushandlung, wer die Frage des Dozenten beantworten soll. Sie folgen dabei einem „Exklusionsmodell“: Nacheinander verdeutlichen drei Team-Mitglieder, dass sie nicht antworten wollen und – in unterschiedlich manifester Weise – dass ein anderer die verbale Antwort übernehmen soll. Wie dieses Modell in detail funktioniert und wie die Aktivitäten der Studenten dabei aufeinander bezogen sind, will ich im Folgenden zeigen.

4. Die interaktive Struktur der Sprechpause

Betrachtet man die sequenzielle Entwicklung der Sprechpause, so lassen sich vier unterschiedliche Phasen identifizieren, die als eigenständige Segmente durch unterschiedliche kinesische Turns im eingeführten Verständnis konstituiert werden. Diese Turns können eindeutig personal zugeordnet werden: Es gibt identifizierbare Aktanten, die durch Bewegungen, Positions- und Blickveränderungen die Interaktion in gleicher Weise weiterführen wie mit verbalen Zügen.

Die im Anschluss vorgestellten Ergebnisse der Rekonstruktion der internen Pausenstruktur präsentieren den Befund eines sequenzanalytischen Vorgehens. Wie bei der Analyse verbaler Äußerungen spielt es dabei eine wesentliche Rolle, wie die Beteiligten mit ihren Verhaltensweisen sequenziell aufeinander Bezug nehmen. Am Anfang der Analyse stand dabei – bezogen auf die vom Dozenten etablierte konditionelle Relevanz – folgende Frage: Was kommt als Nächstes und wie ist dieses Verhalten relativ zu der lokalen interaktiven Anforderung zu interpretieren. Da die lokale Anforderung an die Studenten die Verpflichtung darstellt, auf die Frage des Dozenten zu reagieren, wurde ihr gestisch-körperliches Verhalten als Bearbeitung dieser Aufgabe und als Beitrag zur *turn-taking*-Organisation betrachtet. Ihre Verhaltensweisen wurden dabei auch unter dem Aspekt der Fremd- und Selbstwahl als potenzieller nächster Sprecher/Reagierender analysiert sowie hinsichtlich der Anschlussmöglichkeiten, die damit für die nachfolgend reagierenden Teammitglieder verbunden sind. Darüber hinaus spielte auch die Frage der Adressierung eine Rolle.

Eine ganz wesentliche Orientierung für die Segmentierung der Pause war der Kontrast von Statik und Dynamik. Hierbei zeigte sich, dass jeweils ein/e Student/in primär für die Strukturierung verantwortlich war, indem er/sie erkennbar den *turn* hatte und versuchte, ihn an jemand anderen abzugeben. In der Folge ist das in Segment 1 zunächst Özkan

(der als erster deutlich macht, dass er nicht zur Verfügung steht), dann in Segment 2 alle (indem sie kollektiv nicht reagieren), in Segment 3 Anja ((AN), die wie Özkan zeigt, dass sie nicht antworten will und Martin (MA) auffordert) und in Segment 4 Martin, der – Anja „zitierend“ – ebenfalls signalisiert, dass er nicht bereit ist, die Sprecherrolle zu übernehmen. Schließlich ist es Silvia (SI), welche die Frage des Dozenten verbal beantwortet.

Bild 1: Beginn der Pause (v.l.: Silvia, Martin, Özkan, Anja und Ralf, der Dozent)



Am Ende der Äußerung des Dozenten sind alle Studenten körperlich auf ihn orientiert und ihre Blicke auf ihn gerichtet.

Bild 2: Positionsveränderung Özkan



Segment 1 **Özkan** (Dauer: 3 Sekunden): Unmittelbar nach Ende der Äußerung des Dozenten reagiert Özkan mit einer Veränderung seiner Kopfhaltung und Blickrichtung. Özkan war nicht nur letzter Sprecher der Gruppe, sondern wurde zudem von dem Dozenten mit einer „Stichgeste“ als primärer Adressat seiner Ausführungen hervorgehoben. Özkan wendet seinen Kopf aus der bisherigen Ruheposition leicht nach rechts, hält kurz inne, führt die Kopfbewegung dann weiter, wobei er den Blickkontakt mit dem Dozenten löst und auf die vor ihm auf dem Tisch liegenden Unterlagen schaut. Diese Position friert er ein. Er realisiert diese Bewegungen ohne merkliche Veränderung seiner Körper-

position. Sein Körperdisplay signalisiert eingeschränkte Erreichbarkeit. Während dieser Phase verweilen die anderen Anwesenden in einer stabilen Sitzposition.

Gleiches Bild



Segment 2 *alle* (Dauer 6 Sekunden): Diese Phase ist durch eine vollkommene Statik gekennzeichnet, in der alle Beteiligten regungslos in ihrer Sitzposition verharren (diese Phase ist quasi die „Pause in der Gesprächspause“). Außer Özkan sind alle auf den Dozenten orientiert: Sie sind körperlich auf ihn ausgerichtet, ihre Köpfe orientieren sich in seine Richtung und alle blicken ihn an. Während die Orientierung der drei auf den Dozenten offensichtlich ist, geht dessen Blick leicht nach unten gesenkt in Richtung Özkan und Martin. Die Positionen im Einzelnen:

RA: sitzt leicht zurückgelehnt mit übereinander geschlagenen Beinen, rechtes Knie ragt über Tischplatte, linker Unterarm lässig über Lehne von leerem Nachbarstuhl, linke Hand hängt nach unten, rechter Arm vor Unterkörper, auf rechtem Unterschenkel, Blick leicht nach unten unspezifisch in Richtung MA und ÖZ (Display: *entspanntes Abwarten*; „*ihr seid dran*“)

AN: gerader Sitz, Kopf leicht nach rechts gebeugt, beide Hände unter der Tischplatte, Blick auf RA (Display: *ZuhörerIn*)

ÖZ: leicht vorgebeugter Sitz, Kopf leicht nach vorne geneigt, leicht gespreizte Finger, zu einem Dreieck zusammengeführte Hände, rechter Unterarm auf Tischplatte, Blick auf Unterlagen (Display: *eingeschränkte Erreichbarkeit*)

MA: leicht vorgebeugter Sitz; Kopf gerade, beide Unterarme angewinkelt auf Tischplatte, linke Hand ruht auf rechtem Unterarm, Blick auf RA (Display: *Zuhörer*)

SI: leicht zurückgelehnter Sitz, Kopf links geneigt, beide Hände unter Tischplatte, Blick auf RA (Display: *ZuhörerIn*)

Bild 3: Blickwechsel Anja zu Martin



Segment 3 **Anja** (Dauer: 1,5 Sekunden, fließender Übergang zu Segment 4): wendet ihren Blick an dem rechts neben ihr sitzenden Özkan vorbei und blickt Martin an.

Bild 4: Anja Hand zum Kinn



Leicht verzögert hebt sie ihren rechten Unterarm, führt ihre rechte Hand zum Kinn und stützt dieses auf dem Daumen ab, während der Zeigefinger unter der Unterlippe ruht. Nach Özkan zeigt nun auch Anja mit dieser Geste, die sowohl Nachdenken als auch eine gewisse Ratlosigkeit signalisiert, dass sie im Moment die Sprecherinnenrolle nicht übernehmen will. Bezogen auf den Dozenten ist ihr Körperdisplay das der deutlich eingeschränkten Erreichbarkeit, wohingegen ihr Blick zu Martin klaren Aufforderungscharakter besitzt. Die Botschaft lautet: „*Martin, mach du*“.¹⁵

¹⁵ Jefferson/Schenkein (1978) beschreiben verbale Realisierungen dessen, was die Studentin macht, als „passing“, d.h. als Weitergabe einer Reaktionsverpflichtung an einen anderen Teilnehmer. Das Video-beispiel macht deutlich, dass es neben Formen der aktiven Weitergabe, die mehr oder weniger konkrete Adressaten haben (können), auch solche Reaktionen auf initiative Züge gibt, die man eher als „*let it pass*“ beschreiben kann (Özkans Verhalten hat genau diese Qualität). Der Unterschied zwischen den genannten beiden Formen ist primär ein gesprächsorganisatorischer: „*Let it pass*“ ist eine Form zu zeigen, dass man lokal nicht in die Sprecherrolle eintreten will, ohne dazu die Reaktionsverpflichtung an einen anderen Beteiligten weiter zu geben (und Beteiligte tun dies in der Regel *gerade nicht verbal*). Letztlich handelt es sich hierbei um eine Form der Selbstabwahl.

Bild 5: Martin Hand zum Kinn



Segment 4 *Martin* (Dauer: 1,5 Sekunden): Kurz bevor Anja ihre Hand vollständig zum Kinn geführt hat, nimmt *Martin* die linke Hand vom rechten Unterarm und macht diesen dadurch frei. Mit dem frei gewordenen rechten Arm vollzieht er nun die gleiche Geste wie Anja unmittelbar vor ihm: Er hebt den rechten Arm von der Tischplatte ab, führt seine rechte Hand zum Kinn/Mund, wobei er den Ellenbogen auf der Tischplatte aufstützt. Beide sitzen nun in vergleichbar „nachdenklicher“ Weise am Tisch.¹⁶ Bei genauem Hinsehen wird jedoch deutlich, dass Martin den kinesischen Turn Anjas nicht nur einfach wiederholt. Vielmehr fährt er sich mit den Fingern abwechselnd über Kinn, Mund und Nase. Auch bei ihm ist die Botschaft klar: Ich will die Sprecherrolle ebenfalls nicht übernehmen (er hält sich tatsächlich auch kurzzeitig mit der Hand den Mund zu). Im Unterschied zu Anja fehlt bei ihm eine vergleichbare Adressierung eines anderen Gruppenmitglieds via Blick.

5. Die Sprache kehrt zurück

Im Unterschied zu Özkan, Anja und Martin, die sich und den anderen Beteiligten durch unterschiedliche kinesische Turns wechselseitig angezeigt haben, dass sie – im Augenblick – auf die Frage des Dozenten nicht verbal reagieren wollen, war *Silvia* in der verbalen Auszeit die ganze Zeit auf den Dozenten orientiert: Sie hatte ihre Körperposition nicht verändert noch ihren Blick vom Dozenten gelöst. Sie war an der gesprächsorganisatorischen Aushandlung nicht beteiligt, weder, indem sie selbst hierzu etwas beigetragen hätte, noch dass sie von den anderen Gruppenmitgliedern in die Aushandlung einbezogen worden wäre. Durch ihre stabile Orientierung auf den Dozenten hin etabliert sie gewissermaßen einen eigenständigen Interaktionszusammenhang mit ihm, der im Sinne

¹⁶ De Fornel (1992, S. 160) bezeichnet Gesten wie jene von Martin, die als Wiederholung auf Anjas „Hand-zum-Kinn-Geste“ folgen als *return gestures*. Er hebt hervor, dass *return gestures* „not result from fortuitous juxtaposition of one gesture right after the other one. On the contrary, it has become clear that their organization is in different ways sensitive to (and takes into account) the existence of a (...) gesture in the preceding local sequential environment.“ Der Gestus des Nachdenklichen, der – leicht nach vorne gebeugt sitzend – sein Kinn auf die Hand aufstützt, hat im westlichen Kulturkreis eine lange Tradition. So findet sich beispielsweise eine Miniatur von Walther von der Vogelweide in dieser Pose in der Großen Heidelberger Liederhandschrift aus dem 14. Jahrhundert. Ende des 19. Jahrhunderts wurde „Der Denker“ in dieser Pose von Auguste Rodin in Bronze gegossen.

von Kendon (1984, S. 224) treffend als Interaktionsachse („*axis of interaction*“), beschrieben werden kann. Unter Bezug auf dieses Konzept kann man sagen, dass die anderen Gruppenmitglieder als „*non-axial participants*“¹⁷ agieren.

Silvia zeigt nicht, dass sie auf die Frage des Dozenten nicht reagieren will, und sie versucht auch nicht, die Frage des Dozenten an einen anderen Beteiligten weiterzugeben. Da alle ihre Mitstreiter/innen bereits signalisiert haben, dass sie die Turnvakanz nicht beenden wollen, macht es für sie wenig Sinn, doch noch eine/n von ihnen zu einer Antwort zu bewegen. Sie entscheidet sich vielmehr, auf die Frage des Dozenten inhaltlich-verbal zu reagieren.

Durch ihre Körperhaltung und den konstanten Blickkontakt mit dem Dozenten hat sie sich – anders als die anderen Gruppenmitglieder – in gewisser Weise unmittelbarer und konstant zu der konditionellen Relevanz verhalten bzw. sich nicht von ihr abgewandt. Dieser interaktionsstrukturelle Aspekt koinzidiert mit einem weiteren Aspekt, der verdeutlicht, dass es kein Zufall ist, dass gerade mit Silvia wieder die Sprache in die Interaktion kommt: Silvia ist die Drehbuchautorin des Teams und auf Grund dieser Funktion in besonderer Weise zuständig. In der Rolle der Autorin wird sie von Özkan mit dem Hinweis *so wie das Silvia hier auch schon geschrieben hat * hände zittern und so* explizit ausgewiesen.

Ihre verbale Reaktion, mit der sie die Pause beendet, ist eine Nachfrage, die durch einen deutlich verzögerten Einstieg charakterisiert ist (*ähm äh **) und die an den zweiten Teil der Frage des Dozenten anschließt. Das Format ihrer Reaktion zeigt, dass dieser Frage- teil irritierend ist, da er eigentlich von der Annahme ausgeht, Alex sollte/müsste eine Beziehung zu der vor ihm liegenden Leiche haben. Diese Irritation kam bereits in den kinesischen Turns und den Körperdisplays der anderen Teammitglieder zum Ausdruck.

RA: *12*

SI: ähm äh * muss er denn in einem verhältnis

RA: |ne:" | überhaupt |nicht| * nein nein es geht

SI: |stehen| |nö: |

RA: nur um die um klarheit für diese geschichte

SI: mhm **

Parallel zu Silvias Nachfrage löst Özkan seine linke Hand aus der rechten, hebt den linken Arm an und führt seine linke Hand nun ebenfalls zum Kinn.

¹⁷ Kendon hat diesen Begriff von Watson/Potter (1962) übernommen.

Bild 6: Antwort Silvia; Özkan Hand zum Kinn



Als der Dozent auf Silvias Nachfrage zu antworten beginnt, verändert auch sie ihre bislang stabile Körperposition: Simultan zu seinem Hinweis *nein nein es geht nur um die um klarheit für diese geschichte* löst sie ihre Unterarme und Hände aus der Schoßlage, lehnt sich etwas zurück, kommt dann näher an den Tisch, stützt beide Ellenbogen auf die Tischplatte auf und legt – als sie mit *mhm* auf die Ausführungen des Dozenten reagiert – ihr Kinn auf ihre Handrücken, wobei sie weiterhin Blickkontakt mit dem Dozenten hält. Es ist vor allem dieser konstante Blickkontakt, der in Kombination mit dem leichten Nach-vorne-Rücken und dem auf beiden Handrücken aufgestützten Kinn ein Körperdisplay realisiert, das sich auf den Begriff gebracht als „konzentriertes Zuhören“ beschreiben lässt: Silvia positioniert sich dadurch auch als primäre Adressatin der weiteren Ausführungen des Dozenten.¹⁸

Bild 7: Silvia Kinn auf Hände aufgelegt



Auch Anja hat parallel zur Antwort des Dozenten mit diesem wieder Blickkontakt aufgenommen, so dass nun alle Teammitglieder in einer vergleichbaren Körperhaltung synchronisiert scheinen: Sie sitzen mit aufgestütztem Kinn bzw. Hand (Hände) am Kinn am Tisch und sind auf den Dozenten orientiert. Ein wesentlicher Unterschied besteht jedoch – wie eben beschrieben – in Silvias Positionierung (ihrer Selbstwahl) als primäre Adressatin. In Übereinstimmung mit dieser Markierung ist sie es im Folgenden auch, die wei-

¹⁸ Zur Blickorganisation bei der Markierung von Beteiligungsweise siehe Goodwin (1981).

ter in der Sprecherinnenrolle bleibt: *also äh quasi wie man das darstellt dass er * äh nu“r damit eigentlich probleme hat * jetzt wirklich das skalpell ** rein ** zukriegen.*

Nachdem die Frage, wer die konditionelle Relevanz bearbeitet, geklärt ist, kehren alle Teammitglieder zu ihrer Ausgangsposition vor der Pause zurück: Sie orientieren sich wieder auf den Dozenten, blicken ihn an und verdeutlichen dadurch, dass sie nun wieder zur Verfügung stehen.

6. Analysebefund

Die zurückliegende Analyse folgt den Arbeiten von Scheflen (1964) und Kendon (1970, 1973, 1990), welche die Bedeutung und Implikationen von körperlichen Verhaltensweisen für die Organisation von Gruppenprozessen (das Anzeigen von Beziehungen innerhalb Gruppen sowie der Symbolisierung von Gruppengrenzen) herausgearbeitet haben. Da sich die Analyse auf die sequenzielle Organisation körperlichen Verhaltens im Kontext einer hochgestuften konditionellen Relevanz konzentrierte, konnten zudem die Beziehungen einzelner kinesischer Turns in ihrer Bedeutung für die Bearbeitung der Turnvakanz identifiziert werden. Darüber hinaus war es möglich, die pragmatischen Implikationen einzelner kinesischer Turns zu rekonstruieren. So wurde deutlich, welchen spezifischen Beitrag sie zur Bearbeitung des aktuellen interaktiven Problems leisten, mit dem die Studenten beschäftigt sind.

Zusammenfassend ergibt die Analyse, dass die Studentengruppe in der zwölf Sekunde dauernden Pause zwei unterschiedliche interaktive Zusammenhänge konstituiert: Der eine Zusammenhang wird gebildet durch das körperliche Zusammenspiel von Özkan, Anja und Martin. Der andere besteht – initiiert durch Silvia – zwischen ihr und dem Dozenten.

Das, was auf den ersten Blick wie eine Zersplitterung der Gruppenperspektive aussehen mag, entpuppt sich bei genauem Hinsehen als Ergebnis einer kollektiven Orientierung an der durch die Funktionsrolle gestifteten *primären Zuständigkeit* der Drehbuchautorin: Silvia selbst symbolisiert ihre Zuständigkeit und die Restgruppe behandelt – durch die gestisch-körperliche Organisation eines kollektiven Verzichts auf die Übernahme der Sprecherrolle – Silvia als erste Anwärterin für die Beantwortung der Dozentenfrage.

Das dominante interaktive Problem, das von den Studenten in der Pause durch unterschiedliche kinesische Turns bearbeitet wird, ist das gesprächsorganisatorische Management der Frage, wer die durch den Dozenten gesetzte konditionelle Relevanz bearbeiten soll. Neben diesen gesprächsorganisatorischen Implikationen besitzt das Verhalten der Studenten darüber hinaus einen deutlichen Inszenierungscharakter. Özkan, Anja und Martin symbolisieren ihre innere Befindlichkeit, indem sie ihre Unsicherheit darüber zum Ausdruck bringen, dass sie nicht wissen, was der Dozent mit seiner Frage bezweckt. Zum anderen inszenieren Anja und Martin durch den Einsatz der Denkerpose ihre kognitive Anstrengung. Das Konzept „Inszenieren“ fokussiert interaktives Verhalten, bei dem Teilnehmer lokal multimodale Ressourcen komprimiert einsetzen (verbale Äußerungen, Intonation, Gesten, Körperpostur, Blick, Mimik etc.). Sie stellen damit sicher, dass relevante Implikationen ihrer Verhaltensweisen verstehbar werden. Zum Beispiel denken Interaktionsbeteiligte in bestimmten Situationen nicht nur, sondern zei-

gen, dass sie kognitiv okkupiert sind: Sie wenden ihren Blick leicht nach oben und blicken zu einem imaginären Punkt, platzieren eine Hand unter das Kinn und sind für eine gewisse Zeit für andere Beteiligte nicht unmittelbar erreichbar. Inszenierungen werden von den Beteiligten so zur Bearbeitung aktueller lokaler Interaktionsanforderungen eingesetzt, dass sie sich einen Gewinn bei der Verfolgung ihrer Handlungsziele versprechen.¹⁹

Ein methodischer Aspekt der zurückliegenden Analyse betrifft den analytischen Umgang mit Gesten. Gesten (und andere Formen interaktionsrelevanten körperlichen Verhaltens wie etwa Körperpositur) werden in der Regel in ihrer Bedeutung *als redebegleitende Aktivitäten* untersucht. Dies hat den Vorteil, dass ihre konkrete Platzierung im Interaktionsverlauf ebenso wie ihre Dauer bezogen auf das verbale Geschehen präzise erfasst und beschrieben werden können. Gesten erhalten also einen Teil ihrer Bedeutung nicht aus sich selbst heraus, sondern vielfach von dem mit ihnen realisierten verbalen Geschehen. Sie sind in diesem Sinne, obwohl medial eigenständig, in gewisser Hinsicht *verbal mit-definiert*.

Da bei meinem Beispiel die Handlungskonstitution und die Bearbeitung dominanter interaktiver Aufgaben alleine im Medium des gestisch-körperlichen Ausdrucks erfolgten, entfällt die orientierungsleitende Funktion des Gesprochenen für deren Rekonstruktion. Man ist bei der Rekonstruktion der interaktiven Ordnung, der Herausarbeitung des pragmatischen Potenzials einzelner kinesischer Turns und ihres Bezuges zueinander auf die Gestikulation selbst verwiesen.

Ich bin bei der Rekonstruktion des Beispiels der Sequenzierungsvorstellung der Konversationsanalyse gefolgt, wie sie für die Analyse verbaler Äußerungen entwickelt wurde. Dabei habe ich beispielsweise Anjas Geste als *initiativ* und Martins Geste als *reaktiv* behandelt. Bezogen auf die rein zeitliche Ebene ist dies sicher richtig, analytisch jedoch nicht besonders gehaltvoll. Ein gewisser analytischer Gehalt kommt über eine rein technisch-sequenzanalytische Bestimmung erst dann ins Spiel, wenn man Martins Geste als Reaktion im Sinne der Paarkomplettierung der Gesprächsanalyse rekonstruiert. Dann kann man sagen: Anjas Geste etabliert für Martin eine konditionelle Relevanz, auf die er mit einem kinesischen Turn in Form einer modifizierten „Reformulierung“ reagiert.

So naheliegend eine solche analysepraktische Analogisierung von verbalen und kinesischen Turns auch ist, so stellt sich doch ein latentes Unbehagen ein. Dieses führt zur Frage der grundsätzlichen methodischen Gleichbehandlung verbaler und kinesischer Ausdrucksmittel. Einerseits wird dabei ganz grundsätzlich die Zulässigkeit der Analogisierung thematisiert, andererseits deren Reichweite bzw. die dabei erreichbare analytische Präzision und Eindeutigkeit.

Hinsichtlich der grundsätzlichen Übertragbarkeit der Analysemethode und ihrer zentralen Bestandteile (Sequenzierungsmechanismus, Zeitlichkeit und kontextuelle Spezifik als wesentliche Aspekte der Bedeutungskonstitution) scheinen die Unterschiede bei der Analyse beider Materialtypen nicht kategorial, sondern nur gradueller Art zu sein. Hinsichtlich der dabei erreichbaren Präzision und Eindeutigkeit der Aussagen über gestisch-körperliche Verhaltensweisen gestaltet sich dies problematischer: So ist offensichtlich,

¹⁹ Eine detaillierte theoretische und empirische Einführung in das Konzept „Inszenieren“ findet sich in Schmitt (2003).

dass bei kinesischen Turns bei der Produktion und Analyse eine größere Offenheit des Bezuges und des Anschlusses vorliegt, als dies bei verbalen Äußerungen der Fall ist.²⁰

Ein weiteres Problem ist jenes der Interpretation längerer statischer Phasen in der gestisch-körperlichen Interaktion, wenn alle Beteiligten ihre Aktivitäten einfrieren. Als was und wie kann man solche Phasen körperlicher Inaktivität in Schweigephasen, d.h. die Bewegungslosigkeit in der Stille, interpretieren? Hierfür gibt es idealtypisch zwei Möglichkeiten: Man kann die Bewegungslosigkeit als Symbolisierungsaktivität auffassen und sie als ein Verhalten rekonstruieren, dessen Grundlagen im motivierten Verzicht auf Bewegung besteht. Man kann sie jedoch auch als Symptom konzeptualisieren. In diesem Fall lässt sich nicht viel mehr darüber sagen, als dass die Beteiligten in der aktuellen Situation keine andere Wahl haben als bewegungslos zu verharren.

7. Theoretisierung der gesprächsanalytischen Befunde

Die zwölfsekündige Periode, in der die Gruppe auf verbale Aktivitäten verzichtet, ist der erste Schritt der Studenten, die vom Dozenten etablierte konditionelle Relevanz zu bearbeiten. Die Frage, warum die Gruppe dies in der spezifischen Art und Weise tut, führt zu einer Reinterpretation der oben präsentierten analytischen Ergebnisse.

Die Reinterpretation erfolgt im Hinblick auf zwei unterschiedliche, jedoch eng zusammenhängende theoretische Bezugspunkte. Zum einen werde ich ein *gesprächsrhetorisches Argument* entwickeln, welches das interaktive Verhalten der Studenten während der Pause als Wahl begreift. Zum anderen werde ich diese Wahl in einen *hierarchietheoretischen Rahmen* stellen und das Verhalten als eine Form „kommunikativer Selbstbeschränkung“ charakterisieren.

Gesprächsrhetorische Implikationen

Unter der Voraussetzung, dass die Gruppe nicht hinreichend versteht, worauf der Dozent mit seiner Frage abzielt, was durch Silvias Nachfrage deutlich wird, und unter den eingangs beschriebenen situativen und interaktionsstrukturellen Bedingungen kann man sich fragen, welche Möglichkeiten die Gruppe an dieser Stelle überhaupt hat, ihrer Reaktionsverpflichtung nachzukommen. Diese Frage verweist auf eine gesprächsrhetorische Perspektive.

Die Gesprächsrhetorik fokussiert die Art und Weise, wie Interaktionsbeteiligte unter spezifischen Kontextbedingungen notwendige Konstitutionsanforderungen für ihre eigenen Handlungsziele bearbeiten. Die gesprächsrhetorische Analyse konzentriert sich dabei auf die von den Beteiligten eingesetzten Verfahren und fragt systematisch nach den damit verbundenen Chancen und Risiken.²¹

²⁰ Man kann sich diesen Aspekt verdeutlichen, wenn man bedenkt, dass man – sprichwörtlich – zwar jemanden beim Wort, nicht jedoch bei seiner Geste nehmen kann. Andererseits sind bestimmte gestische Turns nur deswegen Gegenstand juristischer Kommunikation, weil sie die gleichen Implikationen und Konsequenzen haben wie verbale Formulierungen.

²¹ Zum gesprächsrhetorischen Ansatz siehe grundlegend Kallmeyer (1996), zur Darstellung seiner methodischen Arbeitsweise vgl. Kallmeyer/Schmitt (1996).

Aus gesprächsrhetorischer Perspektive ist das Verhalten der Studenten in der Gesprächspause als Selektion einer spezifischen Form der Bearbeitung eines konkreten interaktiven Problems zu begreifen: Es ist nicht nur etwas, was ihnen passiert. Dem Verhalten der Studenten liegt weiterhin eine deutliche Teamorientierung zu Grunde, die als kollektive Orientierung an einem Konzept primärer Zuständigkeit beschrieben worden ist. Diese Orientierung findet ihren interaktiven Ausdruck in der *Selbstabwahl der Nichtzuständigen* und der *Selbstwahl der Zuständigen*. Von den (primär) Nichtzuständigen prescht, solange sich Silvia noch nicht geäußert hat, keine/r vor und thematisiert explizit ihr/sein Nichtverstehen bzw. Verstehensproblem.²² Dadurch, dass Özkan, Anja und Martin sich nacheinander als nicht reaktionswillig positionieren und abwarten, bis die Zuständige reagiert, erfolgt Silvias Antwort zeitlich verzögert.²³ Obwohl es keinerlei Hinweise im körperlichen Verhalten Silvias gibt, die erkennen lassen, ob sie das Selbstabwahl-Procedere der anderen Teammitglieder wahrgenommen hat, kann man davon ausgehen, dass sie – aus den Augenwinkeln – diese Entwicklung genau verfolgen konnte.

Hierarchieindikative Implikationen

Hinreichend geklärt ist jedoch noch nicht, warum die Gruppe auf das gestisch-körperliche Medium und nicht auf das der Verbalität zurückgreift. Hier kommt nun die inszenatorische Qualität ihres Verhaltens ins Spiel. Die Gruppe signalisiert durch ihr Verhalten, dass sie mit der Beantwortung der Dozentenfrage Probleme hat, thematisiert dies jedoch nicht verbal explizit. Unter den gegebenen Situationsbedingungen relativer Fremdheit, dem Dualismus von Gruppen- und Individualorientierung und der latenten Prüfungscharakteristik der Pitching-Situation scheint die Wahl des gestischen Indikationsmodus für ihre Botschaft geeigneter zu sein.

In gewisser Hinsicht kann man das kinesische Verhalten der Studenten in der zwölfsekündigen Sprechpause als eine Form von *delay-Organisation*²⁴ und damit als eine Prä-Sequenz und als Aufschub der von Silvia im Anschluss formulierten Nachfrage verstehen. Es ist Ausdruck einer sensiblen Beziehungsorientierung, die auf explizit interaktionsreflexive Äußerungen und damit verbundene möglichen Beziehungsirritation verzichtet. Erst in einem zweiten Anlauf realisiert Silvia eine solche verbale Äußerung. Ihre Nachfrage ist dabei jedoch erkennbar vorsichtig und weist immer noch äußerungsstrukturelle Aspekte eines *delay-Formats* auf (Nachfrage mit Verzögerung, Pausen und ge-

²² Die Tatsache, dass Özkan erst nachdem Silvia zu sprechen begonnen hat, seine Körperposition verändert und nun ebenfalls seine Hand zum Kinn führt, zeigt als lokales Phänomen, dass sich mit ihrer Reaktion die Situation insgesamt verändert.

²³ Gemessen an der von Jefferson (1989) beschriebenen systematischen Länge von einer Sekunde bei (unproblematischen) Turnwechsel ist die analysierte Schweigephase extrem lang. Dies ist ein Hinweis darauf, unterschiedliche Kategorien von Pausen in Abhängigkeit dominanter Situationsmerkmale und interaktionsstruktureller Bedingungen zu unterscheiden.

²⁴ Heikle, unangenehme oder kritische Sachverhalte werden in der Regel nicht glatt und zielstrebig formuliert, sondern verzögert: Pausen, Modalisierungen, Wortsuche und Reformulierungen etc. schieben die Äußerungsvollendung hinaus. Unter *delay-Organisation* werden Äußerungsrealisierungen verstanden, die als gemeinsames formales Merkmal Aspekte von „*dispreferred turn shapes*“ besitzen, wie sie beispielsweise Pomerantz (1984) im Kontext ihrer Analysen zu Zustimmung und Zurückweisungen von Bewertungen beschrieben hat.

füllten Pausen). Man hat es hier also mit einer Kontinuität formaler Strukturaspekte über einen Wechsel der Modalitätsebenen hinweg zu tun.

Unter Bezug auf das in Schmitt/Heidtmann (2002a) entwickelte Modell der „interaktiven Konstitution von Hierarchie“ kann das Verhalten der Gruppe auch als spezifische Form kommunikativer Selbstbeschränkung verstanden werden. *Kommunikative Selbstbeschränkung* beschreibt die situationsabhängige und situationssensitive mitarbeiterspezifische Variante von Hierarchieorientierung. Kommunikative Selbstbeschränkung setzt dabei die grundsätzliche Akzeptanz der bestehenden formalen Hierarchie voraus.²⁵ Sie kann sich – wie in dem analysierten Beispiel – in einer spezifischen Art und Weise der Bearbeitung interaktiver Probleme realisieren. Das besondere Problemlösungsverfahren der Studenten (kinesisches und verbales *delay*-Format) reagiert dabei primär auf situative Bedingungen. Dominantes Strukturmerkmal sind die klare hierarchische Struktur der Situation (Lehr-Lern-Diskurs) und die damit zusammenhängenden weit reichenden Implikationen, was die Möglichkeit der interaktiven Teilhabe der Studenten betrifft.

Die zwölfsekündige Phase verbaler Abstinenz gehört zur Klasse von Interaktionsausschnitten, welche die Qualität besitzen, relevante strukturelle Eigenschaften der Gesamtsituation, auf kleinstem Raum lokal kondensiert, widerzuspiegeln. Man kann davon sprechen, dass solche Ausschnitte eine mikrostrukturelle Verdichtung der gesamten Situation darstellen.²⁶

Diese Sichtweise deckt sich mit den Ergebnissen einer Studie, die ich gemeinsam mit Daniela Heidtmann im Auftrag der Filmschule zu den organisations- und interaktionsstrukturellen Rahmenbedingungen der Pitching-Situationen durchgeführt habe.²⁷ Diese zeigt die Studenten in einer durchgängig reaktiven Position, die ihnen kaum Raum für selbstbestimmte und selbstorganisierte Darstellungsgelegenheiten lässt. Die Dozenten sind nicht nur inhaltlich, sondern auch in situationskontrollierender und strukturierender Hinsicht die dominanten Agenten.

8. Gesprächsanalyse und multimodale Kommunikation

Aus den vorstehenden Ausführungen ergeben sich weiterführende Fragen, die allgemein das Verhältnis des gesprächsanalytischen Ansatzes und einer Konzeption multimodaler Kommunikation betreffen. Ich möchte mich in diesem Ausblick zumindest noch kurz einigen relevanten Aspekten zuwenden, ohne sie an dieser Stelle angemessen bearbeiten zu können. Diese bestehen in den Implikationen audiovisueller Datengrundlagen und der damit zusammenhängenden Neukonstitution klassischer gesprächsanalytischer Gegenstände (8.1), in der Konstitution neuer Gegenstände (8.2) sowie in der Reflexion klassischer gesprächsanalytischer Konzepte (8.3).

²⁵ Das bedeutet nicht, dass es in bestimmten Situationen nicht doch zur Thematisierung von Erwartungen kommt, die direkt aus der formalen Hierarchie abgeleitet sind.

²⁶ Bei einer solchen Sichtweise ist zu beachten, dass die mikrostrukturelle Verdichtung keine „ontologische“ Qualität des dokumentierten Interaktionsgeschehens selbst ist, sondern sich aus dem theoretisch definierten Erkenntnisfokus des Analysierenden ergibt.

²⁷ Siehe Schmitt/Heidtmann (2002b).

8.1 Neukonstitution klassischer Gegenstände

Die audiovisuelle Datengrundlage tangiert sehr weit gehend die bislang primär auf die verbale Konstitutionsebene fokussierte Gesprächsanalyse. Zum einen verändern sich klassische Gegenstände und Fragestellungen. Dies wird offensichtlich, wenn man – wie zurückliegend – Pausen als relevante Interaktionsphänomene betrachtet. Bei einer ausschließlich verbal basierten Analyse interaktiver Zusammenhänge sind Pausen in der Regel Interaktionspausen. Ihre gesprächsanalytische Behandlung ist beschränkt auf eine Analyse des Vorher und des Nachher. Zu den interaktiven Implikationen der Pause kann dabei aus ihrer internen Struktur selbst nicht viel gesagt werden.

Auf der Grundlage audiovisueller Daten erhalten Pausen einen neuen Status. Sie sind nun nur noch Sprechpausen. Sie werden also als interaktive Phänomene behandelt, die zu Stande kommen, weil Interaktionsbeteiligte sich entscheiden, für eine gewisse Dauer den verbalen Austausch einzustellen. Sie stellen jedoch nicht ihre Interaktion ein, sondern sind vielmehr – vor allem in etwas längeren Pausen – ausgesprochen aktiv. Da Videoaufzeichnungen sichtbar machen, was in Pausen an interaktiv Relevantem geschieht, kommen dieses Geschehen und seine sequenzielle Struktur zur bisherigen „Rand-Analyse“ hinzu. Dabei wird deutlich, dass beispielsweise in den Pausen die *turn taking machinery* weiter läuft, zwar nicht hörbar, aber folgenreich und Struktur prägend.

Zu solch einer Einsicht gelangt man jedoch nur, wenn man auf der Grundlage von Videoaufnahmen beispielsweise den Umgang mit konditionellen Relevanzen während längerer Pausen untersucht. Dann wird deutlich, dass es nicht nur eine Orientierung der Beteiligten an „Redegelegenheiten als kostbarem Gut“ gibt, das geregelt verwaltet werden muss (beispielsweise über Strukturen des *turn taking*). Es wird vielmehr offensichtlich, dass auch der entgegen gesetzte Fall von „Redegelegenheiten als gefährlichem oder unangenehmem Gut“ existiert, bei dem Interaktionsbeteiligte unter den gegebenen lokalen Bedingungen anzeigen, dass sie den Turn nicht übernehmen wollen.

Videoaufzeichnungen ermöglichen so beispielsweise die detaillierte Analyse der Art und Weise, wie von den Beteiligten in sozial akzeptierter und strukturierter Weise *accountable* gemacht wird, dass sie sich nicht als Sprecher etablieren wollen. Und vielleicht steht am Ende solcher Untersuchungen eine Art „*simplest systematics for not taking verbal turns in conversation*“.²⁸ Dann wird Dank der audiovisuellen Grundlagen Folgendes deutlich: Nicht nur die Etablierung als Sprecher ist eine Leistung, die in interaktiver Abstimmung mit den anderen Beteiligten und mittels sequenzieller Einpassung in die lokalen Interaktionsstrukturen realisiert werden muss. Auch die situationsangemessene, sequenziell eingepasste Darstellung des Verzichtes, die Sprecherrolle zu übernehmen, ist eine interaktive Leistung, zu deren Realisierung – wie im analysierten Beispiel – bestimmte Verfahren eingesetzt werden, die es zu beschreiben gilt.

Dies wäre ein möglicher Ansatzpunkt, die bestehende Systematik der Verteilung von Sprecher- und Hörerrollen unter einer multimodalen Perspektive zu einer Systematik der Organisation von Beteiligungsweisen weiter zu entwickeln. Hier könnte man Goffman nicht nur im Sinne Levinsons „*on a proper footing*“, sondern auf eine umfassende multimodale Basis stellen. Somit wäre dies eine Fragestellung an der Schnittstelle von Neu-

²⁸ In Anlehnung an Sacks/Schegloff/Jefferson (1974), einen der wirkungsgeschichtlich zentralen Aufsätze der Konversationsanalyse.

konstitution etablierter Gegenstände und der Konstitution neuer Gegenstände der Gesprächsanalyse.

8.2 Konstitution neuer Gegenstände

Neben der Neukonstitution etablierter Untersuchungsfragen der Gesprächsanalyse treten neue Gegenstände und Fragestellungen hinzu, die an das Medium audio-visueller Repräsentation gebunden sind. Ein solcher Aspekt ist beispielsweise die *kognitive Arbeit und deren inszenatorische Darstellung für die Organisation kooperativer Arbeitsbeziehungen*. Kognitive Arbeit ist ein gleichermaßen interessantes wie methodisch anspruchsvolles Phänomen. Es kann zwar auch auf der Grundlage audiovisueller Daten nicht unmittelbar analysiert werden, wird aber in seiner phänomenologischen Gestalt und seinen multimodalen Ausdrucksformen beschreibbar. Dieser Aspekt lässt sich beispielsweise in Situationen dokumentieren und untersuchen, in denen bestimmte Beteiligte auf der Grundlage einer „*professional vision*“²⁹ das Interaktionsgeschehen systematisch beobachten und dabei über unterschiedliche Displayformen gleichzeitig ihre Verarbeitung des Beobachteten anzeigen.³⁰

Gleiches gilt allgemein für Wahrnehmungs-, Verarbeitungs- und Beteiligungsleistungen von Beteiligten insgesamt. Diese tauchen in den Transkripten bislang primär als Rückmeldeaktivitäten und damit als lokale verbale Reaktionen auf. Auf der Grundlage von Videoaufzeichnungen werden solche Aktivitäten in ihren unterschiedlichen Formen und Funktionen deutlich. Audiovisuelle Grundlagen rücken so zwangsläufig Wahrnehmungsprozesse und deren interaktive Struktur in den Vordergrund.

Als letztes Beispiel für die Neukonstitution von Gegenständen möchte ich auf die Frage nach der interaktiven Etablierung, Aufrechterhaltung und Veränderung von Interaktionsräumen verweisen. Damit ist gemeint, wie Interaktionsbeteiligte die räumlich-situativen Gegebenheiten für ihre Interaktion nutzen und dabei temporär persönliche Territorien etablieren und eine Weile besetzt halten. Die unmittelbare Relevanz solcher Fragen zeigt sich beispielsweise in den Pitching-Aufnahmen: Hier wird deutlich, dass der Platz im unmittelbaren Kontaktbereich zu den Dozenten von den dort sitzenden Studenten eine permanente Verortungsleistung abverlangt: Während die drei anderen Teammitglieder durch ihre Sitzordnung automatisch enger zusammengehören, zeigen die an der Schnittstelle sitzenden Studenten ihre Zugehörigkeit erkennbar dadurch an, dass sie sich räumlich von den Dozenten distanzieren und sich körperlich stark auf die anderen Teammitglieder orientieren. Auffallend dabei ist auch, dass die Studenten in dieser Position sich erkennbar weniger an der gemeinsamen Kooperation beteiligen als die anderen.

8.3 Reflexion klassischer gesprächsanalytischer Konzepte: Das Beispiel *overlap*

Ein letzter Aspekt, auf den ich in diesem kurzen Ausblick zu sprechen kommen möchte, ist die Reflexion etablierter gesprächsanalytischer Vorstellungen und Konzepte, die für

²⁹ Goodwin (1994).

³⁰ Solche expliziten und konzentrierten Monitoring-Aktivitäten sind in den Set-Aufnahmen enthalten: Hier kann man die Regisseure sehen, wie sie Proben der Schauspieler konzentriert verfolgen und im Hinblick auf sehr unterschiedliche Bezugspunkte verarbeiten.

die Analyse der verbalen Seite entwickelt worden sind und sich als ausgesprochen produktiv erwiesen haben. Auch hier will ich mich exemplarisch auf einen Punkt beschränken und an Hand des Konzepts „*overlap*“ zeigen, wie sich die veränderte Datengrundlage und eine konsequent multimodale Analyseperspektive auf dieses Konzept auswirken.

Overlap ist definiert als eine Zeitspanne, in der die grundlegende Orientierung „*overwhelmingly, one party talks at a time*“ (Sacks/Schegloff/Jefferson (1974, S. 706)) außer Kraft gesetzt ist und zumindest zwei oder mehr Interaktionsbeteiligte gleichzeitig sprechen.³¹ Allgemeiner formuliert ist es ein Konzept zur interaktionstheoretischen Beschreibung und konkreten Analyse von simultanen verbalen Aktivitäten. Setzt man an dieser allgemeinen Vorstellung der Simultaneität an, verlässt jedoch nun die Grundlagen einer ausschließlich auf den verbalen Austausch bezogenen Konzeption und überträgt diese auf ein Modell multimodaler Interaktion, für welche die Gleichzeitigkeit und die analytische und theoretische Gleichwertigkeit unterschiedlicher Ausdrucksebenen konstitutiv ist, dann ergibt sich folgendes Bild: Aus einer multimodalen Perspektive ist die gesamte Interaktion ein einziger *overlap*. Zu keinem Zeitpunkt agiert aus multimodaler Sicht ein einzelner Interaktionsbeteiligter alleine, ohne dass simultan nicht andere Beteiligte für den gemeinsamen Interaktionszusammenhang relevante Verhaltensweisen produzieren würden.

In Bezug auf die klassische *overlap*-Vorstellung stellt sich nun die methodologisch interessante Frage, wie sich das verbal definierte Konzept zur multimodalen Vorstellung von Interaktion verhält. Ist es sinnvoll, an dem verbalen *overlap*-Konzept festzuhalten und für die beobachtbare Simultaneität auf anderen Ausdrucksebenen neue Konzepte zu formulieren? Oder ist es sinnvoll, dem Aspekt der Simultaneität so zu begegnen, dass er in einer multimodalen Variante neu formuliert wird? Wie könnte eine solche Neukonzeption dann aussehen?³²

Die Auswirkungen der visuellen Revolution auf die Entwicklung der Gesprächsanalyse und deren etablierte Konzepte betrifft ganz grundlegend auch die Kategorien, die wir für die Beschreibung zentraler Beteiligungsweisen von Sprecher und Hörer zur Verfügung haben. Auch die von Goffman (1979) in seiner *footing*-Konzeption vorgeschlagene Differenzierung und deren Weiterentwicklung von Levinson (1988) sind davon betroffen und folglich im Lichte des nun auch Sichtbaren zu reflektieren.

Ich selbst bin im Augenblick weit davon entfernt, alle Implikationen zu überblicken, die mit einer konsequent multimodalen Analyseperspektive als Reaktion auf die audiovisuelle Qualität der empirischen Grundlagen verbunden sind. Ich sehe jedoch in der Tatsache, dass Analysen von Interaktionsereignissen immer häufiger auf der Grundlage audiovisueller Daten erfolgen, sowohl eine große Herausforderung als auch eine große Chance für die Weiterentwicklung der gesprächsanalytischen Theorie, Methode und Methodologie.

³¹ *Overlaps* sind konstitutiver Bestandteil des *turn-taking*-Systems. Die Systematik des interaktiven Umgangs mit solchen eher kurzen Phasen, in denen mehr als ein Beteiligter spricht, sind bereits in Sacks/Schegloff/Jefferson (1974, S. 706 ff.) skizziert und später beispielsweise in Jefferson (1984) detailliert beschrieben worden. Schegloff hat sich in späteren Arbeiten (2000 und 2001) systematisch mit der interaktiven Struktur von *overlaps* beschäftigt, um der Kritik zu begegnen, die empirische Existenz von *overlaps* sei ein Argument gegen die zentrale *turn-taking*-Annahme von „*one party talks at a time*“.

³² Dieser Aspekt wird eingehend behandelt in Schmitt (i.V. a).

Die multimodale Herausforderung und die damit verbundenen Implikationen gelten jedoch nicht exklusiv für die Gesprächsanalyse, sondern letztlich für alle rekonstruktiven sprachwissenschaftlichen Ansätze, soweit sie auf Gesprächsaufzeichnungen als empirische Grundlage zurückgreifen. Betrachtet man beispielsweise im Kontext der Soziolinguistik eines der zentralen und forschungspraktisch produktivsten Konzepte, nämlich *code-switching*, dann zeigt sich auch hier die Notwendigkeit der multimodal motivierten Reflexion dieses – schon qua Begrifflichkeit eindeutig verbal definierten – Konzeptes.

Multimodale empirische Analysen auf der Grundlage von Videoaufzeichnungen ermöglichen es, systematisch zu klären, ob der Wechsel von einer Sprache zur anderen oder einer Varietät zur anderen tatsächlich ein rein verbales (sprachliches) Phänomen ist. In detaillierten Untersuchungen könnte so rekonstruiert werden, ob sich im Kontext von interaktiv bedeutungsvoller sprachlicher Variation auch noch andere – zwar nicht hörbare, jedoch sichtbare – Dinge im Verhalten der Interaktionsbeteiligten verändern. Gäbe es empirische Evidenz für einen solchen Zusammenhang, wäre der Wechsel auf der bislang untersuchten verbalen Ebene nur ein Aspekt eines umfassenden, multimodalen Phänomens.

Erst wenn auf der Grundlage von Videoaufzeichnungen empirisch hinreichend geklärt ist, dass es sich bei *code-switching* um ein rein sprachliches Phänomen und nicht um einen komplexen multimodalen Vorgang handelt, kann dieses etablierte Konzept in seiner jetzigen Ausarbeitung bestehen bleiben. Sollte sich jedoch zeigen, dass der verbale Wechsel einhergeht mit signifikanten Veränderungen auf anderen Modalitätsebenen, wäre als Konsequenz zu überlegen, wie das Konzept im Hinblick auf die komplexen, multimodalen Eigenschaften von Interaktion – in seinem momentan ausschließlich auf Verbalität reduzierten und damit restriktiven Charakter – modifiziert werden könnte.

9. Zusammenfassung

In diesem Beitrag habe ich am Beispiel einer längeren Gesprächspause gezeigt, wie die gesprächsanalytische Gegenstandskonstitution durch die zur Verfügung stehenden technischen Möglichkeiten beeinflusst wird. Dabei wurde deutlich, dass die Medialität der Daten (hier Videoaufzeichnungen) die Analyse neuer Phänomene ermöglicht bzw. zur Neukonstitution bereits etablierter Untersuchungsgegenstände führt.

Gesprächspausen werden so nicht nur in ihrem Status als konstitutive Interaktionsphänomene deutlich, es zeigt sich vielmehr, dass sie eine sequenzielle Struktur und eine klar erkennbare interaktive Ordnung haben können. Darüber hinaus wird es möglich, das Verhalten der Beteiligten in Analogie zum grundlegenden Sequenzierungsmechanismus der Analyse verbaler Äußerungen in ihrem Status als initiierende und reagierende Paarteile zu beschreiben.

Auf Grund dieser Einsichten ist ein zu enger Kommunikationsbegriff, der auf dem Primat des Gesprochenen gegenüber anderen medialen Ausdrucksformen basiert, zu Gunsten einer multimodalen Konzeption von Interaktion zu erweitern. Deren Aufgabe ist es, für die Analyse einzelner Modalitätsebenen adäquate Analyseverfahren zu entwickeln und dabei zu prüfen, inwieweit sich die für die Analyse verbaler Äußerungen entwickelten Konzepte übertragen lassen. In methodischer Hinsicht scheint mir dabei eine syste-

matische Reflexion gegenstandsadäquater Analyseverfahren wichtig, die nicht vor-schnell von der fraglosen Übertragbarkeit von Prinzipien der Analyse verbaler Äußerungen auf körperliches Verhalten ausgeht.

Wie groß das Ausmaß der Übertragbarkeit letztlich sein wird, ist eine offene Frage und Gegenstand weiterer systematischer Forschung.

10. Anhang: Liste der verwendeten Transkriptionszeichen

SA:	Kennzeichnung des Sprechers/der Sprecherin
K	Sigle für einen Kommentar bezogen auf eine/n Sprecher/in
SB: ja aber	simultan gesprochene Äußerungen stehen
RE: nein nie	untereinander
*	kurze Pause
**	etwas längere Pause (bis max. 1 Sekunde)
3, 5	längere Pause mit Zeitangabe in Sekunden
=	Verschleifung eines Lautes oder mehrerer Laute zwischen Wörtern (z.B. sa=mer für sagen wir)
/	Wort- oder Konstruktionsabbruch
(...)	unverständliche Sequenz
+	sehr schneller Anschluss
↑	steigende Intonation (z.B. kommst du mit↑)
↓	fallende Intonation (z.B. jetzt stimmt es↓)
-	schwebende Intonation (z.B. ich sehe hier-)
“	auffällige Betonung (z.B. aber ge“rn)
:	auffällige Dehnung (z.B. ich war so: fertig)
<immer ich>	langsamer (relativ zum Kontext)
→immerhin←	schneller (relativ zum Kontext)
>vielleicht<	leiser (relativ zum Kontext)
<manchmal>	lauter (relativ zum Kontext)
SB: ach so:	
K IRONISCH	Kommentar zur Äußerung (auf der Kommentarzeile)
LACHT, HUSTET	nichtlexikalisierte Äußerungen

Zitate aus den Transkripten werden im Text *kursiv* wiedergegeben.

11. Literatur

- Bergmann, Jörg R. (1982): Schweigephasen im Gespräch. Aspekte ihrer interaktiven Organisation. In: Soeffner, Hans-Georg (Hg.): Beiträge zu einer empirischen Sprachsoziologie. Tübingen, S. 143-184.
- Birdwhistell, Ray, L. (1952): Introduction to Kinesics. Louisville.
- Birdwhistell, Ray, L. (1960): Kinesics and communication. In: Carpenter, Edmund/McLuhan, Marshall (eds.): Explorations in communication. Boston, S. 54-64.
- Birdwhistell, Ray, L. (1970): Kinesics and context. Essays on body-motion communication. Philadelphia.
- Birdwhistell, Ray, L. (1984): Kinesik. In: Scherer, Klaus, R./Wallbot, Harald, G. (Hg.): Nonverbale Kommunikation: Forschungsberichte zum Interaktionsverhalten. Weinheim (2. Auflage), S. 192-202.
- Butterworth, Brian (1975): Hesitation and semantic planning in speech. In: Journal of Psycholinguistic Research 4, No. 1, S. 75-87.
- Dausendschön-Gay, Ulrich/Krafft, Ulrich (2002): Text und Körpergesten. In: Psychotherapie und Sozialwissenschaft. Zeitschrift für Qualitative Forschung 1/4. Themenheft: Sprechen vom Körper – Sprechen mit dem Körper, S. 30-60.
- Fornel, Michel de (1991): De la pertinence du geste dans les séquences de réparation et d'interruption. In: Conein, Bernard/ Fornel, Michel de /Quéré, Louis (eds.): Les formes de la conversation. Vol 2. Paris, S. 119-153.
- Fornel, Michel de (1992): The return gesture: Some remarks on context, inference, and iconic gesture. In: Auer, Peter/di Luzio, Aldo (eds.): The contextualization of language. Amsterdam/Philadelphia, S. 159-176.
- Goffman, Erving (1961): Encounters. Two studies in the sociology of interaction. Indianapolis.
- Goffman, Erving (1979): Footing. In: Semiotica 25, S. 1-29.
- Goodwin, Charles (1981): Conversational organization. Interaction between speakers and hearers. New York/London.
- Goodwin, Charles (1986): Gestures as a resource for the organization of mutual orientation. In: Semiotica 62 1/2, S. 29-49.
- Goodwin, Charles (1994): Professional vision. In: American Anthropologist 96 (3), S. 606-633.
- Goodwin, Charles (1996): Transparent vision. In: Ochs, Elinor/Schegloff, Emanuel A./Thompson, Sandra A. (eds.): Interaction and Grammar. Cambridge, S. 370-404.
- Goodwin, Marjorie Harness/Goodwin, Charles (1986): Gesture and co-participation in the activity of searching for a word. In: Semiotica 62-1/2, S. 51-75.
- Heath, Christian (1982): The display 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 (1992): Gesture's discreet tasks: Multiple relevancies in visual conduct and in the contextualisation of language. In: Auer, Peter/di Luzio, Aldo (eds.): The Contextualization of Language. Amsterdam, S. 101-127.
- Jefferson, Gail (1984): Notes on some orderliness of overlap onset. In: D'Urso, Valentina/Leonardi, Paolo (eds.): Discourse analysis and natural rhetorics. Padova, S. 11-38.
- Jefferson, Gail (1989): Preliminary notes on a possible metric which provides for a 'standard maximum' silence of approximately one second in conversation. In: Roger, Derek/Bull, Peter (eds.): Conversation: An interdisciplinary perspective. Clevedon, S. 166-196.

- Jefferson, Gail/Schenkein, Jim (1978): Some sequential negotiations in conversation. Unexpanded and expanded versions of projected action sequences. In: Schenkein, Jim (ed.): *Studies in the organization of conversational interaction*. New York, S. 155-172.
- Kallmeyer, Werner (1996): Einleitung. In: Ders. (Hg.): *Gesprächsrhetorik. Rhetorische Verfahren im Gesprächsprozeß*. Tübingen (Studien zur deutschen Sprache 4), S. 7-18.
- Kallmeyer, Werner/Schmitt, Reinhold (1996): Forcieren oder: Die verschärfte Gangart. Zur Analyse von Kooperationsformen im Gespräch. In: Kallmeyer, Werner (Hg.), S. 20-118.
- Kendon, Adam (1970): Movement coordination in social interaction: Some examples described. In: *Acta Psychologica* 32, S. 100-25.
- Kendon, Adam (1972): Some relationships between body motion and speech. In: Siegman, Aron Wolfe/Pope, Benjamin (eds.): *Studies in dyadic communication*. New York, S. 177-210.
- Kendon, Adam (1973): The role of visible behavior in the organization of social interaction. In: Cranach, M. v./Vine, L. (eds.): *Social communication and movement*. London, S. 29-74.
- Kendon, Adam (1980): Gesticulation and speech: Two aspects of the process of utterance. In Key, Marie Ritchie (ed.): *The relationship of verbal and nonverbal communication*. The Hague, S. 207-228.
- Kendon, Adam (1984): Die Rolle sichtbaren Verhaltens in der Organisation sozialer Interaktion. In: Scherer, Klaus R./Wallbott, Harald G. (Hg.): *Nonverbale Kommunikation: Forschungsberichte zum Interaktionsverhalten*. Weinheim, S. 202-235.
- Kendon, Adam (1990): *Conducting interaction. Patterns of behavior in focuses encounters*. (Studies in Interactional Sociolinguistics 7). Cambridge.
- Kendon, Adam (1994): Do gestures communicate? A Review. In: *Research on Language and Social Interaction* 27(3), S. 175-200.
- McNeill, David (1992). *Hand and mind: What gestures reveal about thought*. Chicago.
- McNeill, David (2000) (ed.). *Language and Gesture: Window into Thought and Action*. Cambridge.
- Levinson, Stephen C. (1988): Putting linguistic on a proper footing: Explorations in Goffman's concept of participation. In: Drew, Paul/Wootton, Anthony (eds.): *Erving Goffman: exploring the interaction order*. Abingdon, S. 161-227.
- Müller, Cornelia (1998): *Redebegleitende Gesten. Kulturgeschichte – Theorie – Sprachvergleich*. Berlin.
- Müller, Cornelia (2003): On the gestural creation of narrative structure: A case study of a story told in conversation. In: Poggi, Isabella/Rector, Monica/Trigo, Nadine (eds.): *Gestures: Meaning and Use*. Porto, S. 259-265.
- Müller, Cornelia/Paul, Ingwer (1999): Gestikulieren in Sprechpausen. Eine konversationsanalytische Fallstudie. In: Eggert, Hartmut/Golec, Janusz (Hg.): *„wortlos der Sprache mächtig“: Schweigen und Sprechen in Literatur und sprachlicher Kommunikation*. Stuttgart, S. 265-282.
- Pomerantz, Anita (1984): Agreeing and disagreeing with assessments: Some features of preferred/dispreferred turn shapes. In: Atkinson, J. Maxwell/Heritage, John (eds.): *Structures of social action. Studies in conversation analysis*. Cambridge, S. 57-101.
- Rochester, S.R. (1973): The significance of pauses in spontaneous speech. In: *Journal of Psycholinguistic Research* 2 (1), S. 51-81.
- Sacks, Harvey/Schegloff, Emanuel A./Jefferson, Gail (1974): A simplest systematics for the organization of turn taking for conversation. In: *Language* 50 (4), S. 695-735.
- Schefflen, Albert E. (1963): Communication and regulation in psychotherapy. In: *psychiatry* 26, S. 126-136.
- Schefflen, Albert E. (1964): The significance of posture in communication systems. In: *Psychiatry* 27, S. 316-313.

- Schefflen, Albert E. (1972): *Body language and social order: Communication as behavioral control*. Englewood Cliffs.
- Schefflen, Albert E. (1973): *Communicational structure. Analysis of a psychotherapy transaction*. Bloomington.
- Schegloff, Emanuel A. (1972): Sequencing in conversational openings. In: Gumperz, John J./Hymes, Dell (eds.): *Directions of Sociolinguistics. The Ethnography of Communication*. New York/Chicago, S. 346-380.
- Schegloff, Emanuel A. (1984): On some gestures' relation to talk. In: Atkinson, Maxwell J./Heritage, John C. (eds.): *Structures of social action*. Cambridge, S. 266-296.
- Schegloff, Emanuel A. (2001): Accounts of conduct in interaction. Interruption, overlap, and turn-taking. In: Turner, Jonathan H. (ed.): *Handbook of Sociological Theory*. New York, S. 287-321.
- Schegloff, Emanuel A. (2000): Overlapping talk and the organization of turn-taking for conversation. In: *Language in Society* 29, S. 1-63.
- Schegloff, Emanuel A./Sacks, Harvey (1973): Opening up closings. In: *Semiotica* 8, S. 289-327.
- Schmitt, Reinhold (2001): Von der Videoaufzeichnung zum Konzept „Interaktives Führungshandeln“: Methodische Probleme einer inhaltlich orientierten Gesprächsanalyse. In: *Gesprächsforschung – Onlinezeitschrift zur verbalen Interaktion* (ISSN 1617-1837) Ausgabe 2 (2001), S. 141-192 (www.gespraechsforschung-ozs.de).
- Schmitt, Reinhold (2003): Inszenieren. Struktur und Funktion eines gesprächsrhetorischen Verfahrens. In: *Gesprächsforschung – Onlinezeitschrift zur verbalen Interaktion* (ISSN 1617-1837), Ausgabe 4 (2003), S. 186-250 (www.gespraechsforschung-ozs.de).
- Schmitt, Reinhold (2004): Bericht über das 1. Arbeitstreffen „Multimodale Kommunikation“. In: *Sprachreport* 1/2004, S. 31-34.
- Schmitt, Reinhold (i.V. a): Zur multimodalen Struktur von turn taking. Das Beispiel *overlap*. Erscheint in: *Gesprächsforschung – Onlinezeitschrift zur verbalen Interaktion* (ISSN 1617-1837), (www.gespraechsforschung-ozs.de).
- Schmitt, Reinhold (i.V. b): „Silent machinery“. Kinesic turns in a 12-second speech pause. Erscheint in: *Gesture*.
- Schmitt, Reinhold/Heidtmann, Daniela (2002a): Die interaktive Konstitution von Hierarchie in Arbeitsgruppen. In: Becker-Mrotzek, Michael/Fiehler, Reinhard (Hg.): *Unternehmenskommunikation*. Tübingen, S. 179-208.
- Schmitt, Reinhold/Heidtmann, Daniela (2002b): Abschlussbericht der Studie „Pitching in Teams“. Bestandsaufnahme und erste Vorschläge. Ms. Mannheim.
- Schmitt, Reinhold/Heidtmann, Daniela (2003a): „Linguistik goes Hollywood“: Angewandte Gesprächsforschung in der Ausbildung von Filmstudenten. In: *Gesprächsforschung – Online-Zeitschrift zur verbalen Interaktion* (ISSN 1617-1837), Ausgabe 4 (2003), S. 98-121 (www.gespraechsforschung-ozs.de).
- Schmitt, Reinhold/Heidtmann, Daniela (2003b): Angewandte Gesprächsforschung in der Filmbranche. In: *Sprachreport* 3/2003, S. 11-17.
- Seyfeddinipur, Mandana/Kita, Sotaro (2001): Gestures and dysfluencies in speech. In: Cavé, Christian/Guaitella, Isabelle/Santi, Serge (eds.): *Oraligé et Gestualité. Interactions et comportements multimodaux dans la communication*. L'Harmattan, S. 266-270.
- Streeck, Jürgen (1994): Gesture as communication II: The audience as co-author. In: *Research on Language and social Interaction* 27 (3), S. 239-267.
- Streeck, Jürgen (1995): On projection. In: Goody, Esther N. (ed.): *Social intelligence and interaction: expressions and implications of the social bias in human intelligence*. Cambridge, S. 84-110.

Streeck, Jürgen/Hartge, Ulrike (1992): Previews: Gestures at the transition place. In: Auer, Peter/di Lucio, Aldo (Hg.): The contextualization of language. Amsterdam/Philadelphia, S. 135-157.

Tönnies, Ferdinand (1982): Die Tatsache des Willens. Berlin.

Umiker-Sebeok, Jean D. (1980): Silence is golden? The changing role of non-talk in preschool conversation. In: Key, Marie Ritchie (ed.): The relationship of verbal and nonverbal communication. The Hague, S. 295-314.

Watson, J./Potter, R.J. (1962): An analytic unit for the study of interaction. In: Human Relations 15, S. 243-243.

Dr. Reinhold Schmitt
Institut für Deutsche Sprache
Postfach 101621
D-68016 Mannheim
e-mail: reinhold.schmitt@ids-mannheim.de