

4 Kopräsenz und verbale Interaktion: Interaktionsprobleme und ihre Lösung

Wenn immer sich Kopräsenz einstellt und über die Flüchtigkeit einer Augenblicksepisode hinausragt, haben wir es mit einem ausgewachsenen Fall von Interaktion zu tun, d.h. mit der Lösung einer Reihe interaktionskonstitutiver Aufgaben. Viel spricht dafür, dass diese Aufgaben in dem Maß an Relevanz und Profil gewonnen haben, in dem Kopräsenz durch Sprache konfiguriert wird und Sprache entsprechend als Ressource der Interaktion in Erscheinung tritt und von den Beteiligten genutzt wird. Es sind deshalb, bis auf eine Ausnahme, vordringlich Probleme der *verbalen* Interaktion, die wir in diesem Kapitel behandeln wollen – was nicht heißt, dass diese Probleme immer und alle nur durch und mit Sprache bearbeitet und »gelöst« werden.

Die Einsicht in die Probleme der Interaktion verdankt sich vor allem einer von Anfang an fast exklusiv auf *verbale* Interaktion fokussierten Forschungsrichtung: der kalifornischen Konversationsanalyse. Sie hat in gewisser Weise empirisch entdeckt, was man im Anschluss an ein theoretisches Kernkonzept der Systemtheorie die »Unwahrscheinlichkeit der Interaktion« nennen könnte. Entgegen unserer Alltagsintuition, die sich der allgegenwärtigen Teilnahme an einer kaum unterbrochenen Kette von Interaktionsepisoden verdankt, versteht sich der Alltag der Interaktion nicht von selbst. Dass wir einander begrüßen und verabschieden, uns im Gespräch abwechseln und Beiträge zu Themen austauschen, uns und unser Gegenüber dabei durch Selbst- und Fremddarstellungen ein- und zuordnen und den Austausch gesellschaftlich rahmen und situativ verankern, gehört wohl zur selbstverständlichen Alltagsroutine, über die wir weder nachdenken noch grübeln, wenn es nicht zu praktischen Handlungsproblemen kommt (derart, dass man z.B. trotz wiederholter Versuche nicht aus einem Gespräch herauskommt). Aber diese Routine ist nur scheinbar selbstverständlich und natürlich, wenn man damit beginnt, sich die Probleme zu vergegenwärtigen, die damit einhergehen. Es erfordert dafür eine bewusst

kontraintuitive Verfremdung, die erklären hilft, wie an sich unwahrscheinliche Interaktionsvorgänge wahrscheinlich und erwartbar gemacht werden, so dass sie uns im Alltag als solche gar nicht mehr auffallen. Dass die Unwahrscheinlichkeit der Interaktion in jeder Interaktionsepisode wie selbstverständlich überwunden wird, ist also Teil der Evolution kopräsenzbasierter Interaktionssysteme.¹ Damit fangen wir gleich an.

Anschließend gehen wir über zu typischen Problemen der Interaktion. Interaktion lässt sich als alltagspraktische Bearbeitung und Lösung wiederkehrender Anforderungen verstehen. Das fängt mit der Interaktionseröffnung und der Interaktionsbeendigung an, die zu den frühen konversationsanalytischen Gegenständen gehören, und führt nahtlos weiter zum Problem des Sprecherwechsels, dem die Konversationsanalyse bis heute ihr Renommee verdankt. Weniger prominent, aber nicht weniger wichtig sind weitere Themenstellungen, die mit der gesellschaftlichen Rahmung der Interaktion, der Selbst- und Fremddarstellung und der Verankerung der Interaktion im Hier und Jetzt der Anwesenden zu tun haben. Unabhängig davon, ob diese Liste als vollständig betrachtet werden kann oder nicht, zeigt sie eindrucksvoll, wie sich das Axiom der Unwahrscheinlichkeit der Interaktion im Rückgriff auf konversations- und gesprächsanalytische Forschung empirisch fruchtbar machen lässt. Wichtiger für unsere Themenstellung ist aber, dass ein Blick auf diese Aufgaben und ihre empirischen Lösungen vor Augen führt, wie sehr Kopräsenz mit dem Auftauchen von Sprache refiguriert und profiliert wird. Das wollen wir mit einem kursorischen Durchgang durch die wichtigsten Interaktionsaufgaben entwickeln.

Unter den Problemen der Interaktion ist das Reihenfolgeproblem womöglich das prominenteste. Es manifestiert sich in der Aufgabe des Sichabwechselns (Wer kommt als Nächste oder Nächster?) und mag als solches auch schon vor dem Einsatz von Sprache aufgetreten sein. Aber erst mit und durch Sprache wird es zu einem Kardinalproblem der Interaktion. Es ist deshalb vielleicht nicht zufällig gleich am Anfang der konversationsanalytischen Forschung sehr prominent in Erscheinung getreten. Im Gegensatz dazu hat das Problem der Verankerung der Interaktion im Hier und Jetzt der Anwesenden, das wir als Problem der *Situierung* einführen werden, in der Konversations- und Gesprächsanalyse lange ein Schattendasein gefristet. Tatsächlich ist es ein Problem, das gewissermaßen noch vor dem Sprechen-und-Zuhören einsetzt und dessen Bearbeitung und Lösung beim Sprechen-und-Zuhören in vielen Fällen bereits vorausgesetzt wird. Es führt also onto- wie phylogenetisch weiter zurück. Man kann daran viel über basale Formen von Kopräsenz

lernen, die noch nicht oder doch nicht primär auf Sprache verweisen, sondern primär auf Wahrnehmung (Ko-Orientierung) und Bewegung (Ko-Ordination), wie sie in Situationen komplexer Handlungen (Ko-Operation) in der Regel vorausgesetzt werden. Die Situierung ist vor allem in den letzten Jahren mit der (Wieder-)Entdeckung des Interaktionsraums in das Bewusstsein der konversationsanalytischen Forschung getreten. Wir widmen der Aufgabe der Situierung einen eigenen Abschnitt und gehen anschließend auf die aus der Situierung erwachsenden Probleme der verbalen Interaktion ein.

Die Unwahrscheinlichkeit der Interaktion

Kopräsens ist entgegen der Mythen, die mit ihr als direkt-unmittelbarer, gleichsam primordial-natürlicher Sozialform verbunden sind, alles andere als selbstverständlich. Sie ist, wie wir wiederholt betont haben, nicht einfach gegeben, sondern muss kommunikativ hergestellt werden. Sie ist in genau diesem Sinne eine evolutionäre Errungenschaft, die nicht nur für Sprache, sondern allgemein für die Menschwerdung von zentraler Bedeutung ist. Diese allgemeine Annahme soll in diesem Abschnitt konkretisiert werden mithilfe des Axioms der »Unwahrscheinlichkeit der Interaktion«. Es besagt, dass Interaktion entgegen unserer Vertrautheit mit ihrem alltäglichen Vollzug ein höchst problematisches, immerzu vom Abbruch bedrohtes Unterfangen ist. Man muss also erklären können, warum Interaktion trotzdem im Alltag immer wieder gelingt und möglich und wahrscheinlich werden kann. Dieses theoretische Axiom übernehmen wir aus der neueren soziologischen Systemtheorie.² Empirisch orientiert es sich an der Forschungspraxis der Konversations- und Gesprächsanalyse. Dabei muss berücksichtigt werden, dass sich diese Forschungsrichtung zumindest in ihren Anfangsjahren so gut wie ausschließlich und bis heute vornehmlich mit *verbaler* Interaktion beschäftigt hat.³ Die Konversationsanalyse hat die Unwahrscheinlichkeit der Interaktion also wie selbstverständlich am Gegenstand der *sprachlich* geprägten Interaktion entdeckt, und es ist noch die Frage, ob man von dieser Konfiguration von Kopräsens durch Sprache ohne weiteres abstrahieren kann. Das gilt es im Auge zu behalten, wenn wir im Folgenden den einschlägigen Forschungsstand im Hinblick auf die Unwahrscheinlichkeitsannahme rekapitulieren.

Die Analyse von Interaktion, die in verschiedenen Disziplinen, vorrangig aber in der Linguistik und Soziologie, unter verschiedenen Bezeichnungen,

vorrangig als Gesprächs- und Konversationsanalyse, betrieben wird, kann inzwischen auf eine mehr als 50-jährige Tradition zurückblicken, in der sich bei allen Unterschieden und Varianten eine Standardisierung der Vorgehensweise ergeben hat. Diese interaktionsanalytischen Standards sind in den letzten Jahren erfreulicherweise qualitativ wie quantitativ runderneuterten Einführungsliteratur ausführlich beschrieben worden.⁴ Wir können uns deshalb auf einige grundlegende methodologische Überlegungen beschränken, ohne dabei zu technisch werden zu müssen. Wir verteilen diese Überlegungen auf mehrere Abschnitte: Zunächst geht es um grundsätzliche Zugangsweisen vom Feldzugang über die Datenerhebung bis zur Datenaufbereitung. Eine herausragende Bedeutung kommt dabei dem Transkribieren und der Transkription zu: Die Transkription ist als Sekundärdatum ein in seiner Bedeutung kaum überschätzbares Schlüsseldokument konversationsanalytischer Forschung, an dem man viel über die Unwahrscheinlichkeit der Interaktion lernen kann. Weiter wird es um Grundprinzipien des Umgangs mit den erhobenen und aufbereiteten Daten und der Datenauswertung gehen. Diese Grundprinzipien lassen sich im Sinne des Unwahrscheinlichkeitsaxioms reformulieren, wie wir im Anschluss an Niklas Luhmann zeigen wollen. Die Grundprinzipien zwingen zu einer weitgehenden Verfremdung der Alltagsperspektive, die zu einem Bruch mit der Selbstverständlichkeit alltäglich erlebter Kopräsenz führt. Warum ein solcher Bruch zwingend notwendig ist, soll einleitend mit einem kleinen Exkurs illustriert werden.

I.

Am Anfang einer Rede, die der US-amerikanische Schriftsteller David Foster Wallace 2005 vor einer Abschlussklasse am Kenyon College (Ohio) gehalten hat und die mit dem Titel »This is water« überschrieben ist,⁵ steht eine kurze gleichnishafte Geschichte, auf die wir bereits in der Einleitung zu diesem Essay hingewiesen haben. Wir geben diese kurze Geschichte hier wieder, weil sie sehr anschaulich erhellte, dass die »naheliegendsten, allgegenwärtigsten und wichtigsten Dinge im Leben oftmals zu den Dingen gehören, die am schwersten zu bemerken und zu besprechen sind.«⁶

Die Geschichte geht wie folgt:

There are these two young fish swimming along and they happen to meet an older fish swimming the other way, who nods at them and says, »Morning, boys. How's the water?«

And the two young fish swim on for a bit, and then eventually one of them looks over at the other and goes, »What the hell is water?«⁷

Könnte es sein, dass das, was für die Fische das Wasser ist, für uns Menschen als soziale Wesen die Kooperation in Kopräsenssituationen ist, unser soziales Lebenselixier, die soziale Luft, die wir atmen, wenn wir kommunizieren? Und könnte es uns, die wir im Alltag immer wieder in Kopräsenssituationen stecken, so gehen wie den beiden »jungen Fischen« mit dem »Wasser«? Dass wir auf Fragen nach diesem sozialen Zuhause unseres sozialen Lebens ähnlich irritiert reagieren würden: »Was, zum Teufel, ist Kopräsens?« Kopräsens gehört zu unserem Alltag wie das Essen, Trinken, Schlafen und Träumen. Wir sitzen vis-à-vis mit Unbekannten zusammen in der Straßenbahn, treffen unterwegs Leute, die wir grüßen oder denen wir gekonnt aus dem Weg gehen, verbringen Pausen und Mittag- und Abendessen zusammen mit Freunden, Freundinnen, Bekannten, Kolleginnen und Kollegen, mögen vielleicht nicht einmal vor dem nächtlichen Einschlafen allein sein und sind dabei und darüber hinaus immer wieder in Gespräche verstrickt. Wir erleben, wie ein Wort das andere gibt, und dass Reden Silber und Schweigen oft Gold ist. Unser Kopräsensalltag ist voll von Anforderungen der Ko-Orientierung (Wer ist noch da?), der Ko-Ordination (Worauf muss ich achten?) und der Ko-Operation (Was machen wir zusammen?), denen wir in der Regel mehr oder weniger routinisiert gerecht werden, ohne dass wir darüber nachdenken und grübeln und ohne dass wir darüber mit Anderen sprechen oder diskutieren müssten. Im Prinzip läuft dieser Kopräsensalltag weitgehend unbemerkt ab. Wenn wir anfangen, über das nachzudenken, was im Gespräch und anderen sprachlosen Episoden von Kopräsens passiert, ist deshalb in der Regel schon alles entschieden.

So selbstverständlich diese Praxis ist, so unverständlich ist sie freilich auch. Dass Kopräsens auf Wahrnehmungswahrnehmung beruht, also darauf, dass ich z.B. sehen kann, dass ich gesehen werde, und wiederum auch dieses Sehen des Gesehenwerdens sehe (und so fort in unabschließbarer Selbstbezüglichkeit), ist eine im Alltag typischerweise immer schon beantwortete, aber nicht reflektierte Anforderung. Ihre Bearbeitung ist daher nicht darauf angewiesen, dass sie die Bewusstseins- und Verbalisierungsschwelle der Beteiligten überschreitet. Auch müssen die Beteiligten beim Sprechen-und-Zuhören nicht bei laufendem Motor mitvollziehen, wie sie es immer wieder fertigbringen, sich ohne größere Überlappungen abzuwechseln und mehr oder weniger sekundengenau zu wissen, wie lange sie *dran* sind, wenn sie selbst sprechen, und wie sie überhaupt *drankommen* im Gespräch, wenn sie

gerade (nur) zuhören. In allen Problem- und Aufgabenbereichen der Interaktion, die wir noch besprechen werden, verhält es sich ähnlich: Der Vollzug von Interaktion mit und durch die Erledigung interaktionskonstitutiver Anforderungen verläuft in der Regel unterhalb der Aufmerksamkeitsschwelle der Beteiligten und ragt nur dann in die Zone bewusster und thematisierter Reflexion, wenn etwas schiefgeht oder nicht so (reibungslös) klappt, wie man es gewohnt ist, und es z.B. zu mehr oder weniger aufwendigen »Reparaturen« kommt – die dann ihrerseits wieder einer Regel- und Gesetzmäßigkeit folgen, die selbst nicht reflektiert werden muss. Darin ist kein Versäumnis der Interaktionsbeteiligten zu sehen, sondern im Gegenteil eine *Conditio sine qua non* der Interaktion, in der sich anschaulich zeigt, dass Interaktion ein eigen-gesetzlich soziales Geschehen ist, das sich *zwischen* den Beteiligten abspielt. Im Alltag haben wir dafür als Interaktionsbeteiligte unter Handlungsdruck weder ein Bewusstsein noch eine Sprache. Die bekannt technische Terminologie der Konversationsanalyse zur Beschreibung des Sprecherwechsels (»turn taking«) gehört zur Sprache der Beobachter:innen, nicht zur Sprache der Teilnehmer:innen. Gesprächsbeteiligte beobachten den Sprecherwechsel aus praktischen Interessen und für praktische Belange des Drankommens und Dranbleibens, und sie vollziehen ihn auf dieser Grundlage, ohne ihn dafür auf eine Weise zu verstehen, dass sie darüber Auskunft geben könnten, wenn sie gefragt würden. Es ist, wenn überhaupt, ein stillschweigendes Wissen, das nicht von der Anwendung in einer konkreten Situation abgezogen werden kann. Und auch die Reflexion ist nicht von ihrer Anwendung in einer konkreten Situation abzuziehen. Interaktionsbeteiligte neigen deshalb dazu, Interaktion in Termini des Erlebens und Handelns⁸ zu reflektieren, das sich selbst und Anderen in Form von Motiven und Intentionen zugeordnet, zugeschrieben, bewertet und verstanden werden kann. An- und Abwesenheit werden entsprechend in der Regel als gewählt erlebt, Kopräsenz als ein sich daraufhin einstellendes Zusammensein mit Anderen behandelt, die ebenfalls anwesend sein wollen. Die genuin sozialen Regel- und Gesetzmäßigkeiten, die hinter solchen vereinfachenden Mechanismen des Erlebens und Handelns stecken, bleiben entsprechend unbemerkt – und unentdeckt. Erst unter der Bedingung der Handlungsentlastung, wie sie sich mit der Erhebung von Kopräsenzdaten in Form von Audio- und Videoaufzeichnungen von Interaktionsereignissen methodisieren lässt, entsteht die Einsicht in die Unwahrscheinlichkeit der Interaktion. Das wollen wir im Folgenden nachzeichnen.

II.

Der erste Schritt auf dem Weg einer empirisch motivierten Einsicht in die Unwahrscheinlichkeit der Interaktion besteht in der Art und Weise der Erhebung und Aufbereitung von Interaktion. Wenn es eine Gemeinsamkeit der Interaktionsanalyse über alle Schulen und Richtungen hinweg gibt, besteht sie darin, dass die Analyse von Interaktion auf Daten vom Typus der Audio- (und Video-)aufzeichnung⁹ konkreter Interaktionsereignisse angewiesen ist. Diese Feststellung ist in zweierlei Hinsicht aufschlussreich, die wir erläutern wollen, ohne dabei zu »methodisch« zu werden. Zum einen muss man das Votum für Aufzeichnungen vor dem Hintergrund alternativer bzw. komplementärer Zugänge zur Erforschung von Interaktion verstehen, und zum anderen muss man die Aufwertung der hörbaren sprachlichen Erscheinungsformen von Kopräsens nachvollziehen, die mit der Entscheidung für Daten in Form von Audio- (und Video-)aufzeichnungen von Interaktion verbunden ist. Wir gehen zunächst auf die zuletzt genannte Favorisierung *verbaler* Interaktion ein und kommen dann auf die Entscheidung für die Dokumentation von Interaktion in Form der Aufzeichnung zurück.

Das Interesse für genuine Daten, die Aufschluss über kopräsensbasierte Interaktion geben könnten, ist *zum einen* eng mit der Möglichkeit der Audioaufzeichnung von Gesprächen verbunden, also der Dokumentation von gesprochener Sprache durch Audiorekorder. Man kann auch sagen: Mit dem Sprechen-und-Zuhören, mit der Konfiguration von Kopräsens durch Sprache, werden Kopräsenssituationen in ihren akustischen Manifestationen auf eine neuartige Weise dokumentierbar, was sich dann im Nachhinein als Schlüssel für die Einsicht in die Unwahrscheinlichkeit der Interaktion und die Dekonstruktion des Kopräsensmythos erwiesen hat. Zwar gab es mit der Kamera auch schon früher die Möglichkeit der Dokumentation von Kopräsenssituationen.¹⁰ Aber die daraus resultierenden Daten haben lange vorrangig der Begleitung und Unterstützung der Aufzeichnung des Gesprochen-Gehörten gedient.¹¹ Der beste Beleg dafür ist das Konzept der »nonverbalen Kommunikation«, das sich wohl nicht trotz, sondern gerade wegen seiner Orientierung an der »Verbalität« als reine Ausschlussdefinition für all das verbreitet und durchgesetzt hat, was unter Kopräsensbedingungen neben dem Gesprochen-Gehörten noch wahrzunehmen ist. Es kommt hinzu, dass einer der innovativsten und erfolgreichsten Ansätze zur Interaktionsanalyse, die Konversationsanalyse, mit dem Telefongespräch eine Konfiguration von Kopräsens zum Gegenstand gemacht hat, die ausschließlich auf Hörbarkeit

beruht. Kopräsenz wird in diesem Fall dominant sprachlich manifest, so dass eine auf die Aufzeichnung des Hörbaren reduzierte Erhebungstechnik die sprachliche Konfiguration von Kopräsenz an einem besonders eindrücklichen Fall erfassen kann – wenngleich dabei die Besonderheit dieser Konfiguration von Kopräsenz noch zu reflektieren wäre (worauf wir noch zurückkommen).¹² Der Rückblick zeigt aber, wie sehr der Beginn der Interaktionsforschung durch die Beschäftigung und Dokumentation *sprachlicher* Interaktion geprägt ist. Die darin zum Ausdruck kommende Gleichsetzung von Kopräsenz mit verbaler Interaktion ist nicht trivial, weil sie die Konzentration vor allem auf solche Probleme der Interaktion begünstigt hat, die besonders stark durch Sprache geprägt sind – wie das schon erwähnte Problem des Sprecherwechsels, das zudem am Telefon in seiner sprachlichen Konstituiertheit markant hervortritt.

Zum anderen ist auch schon die oben thematisierte Entscheidung für die Dokumentation von Interaktion in Form der audiovisuellen Aufzeichnung nicht trivial. So gibt es bekanntermaßen auch andere Zugänge zur Interaktion, die auf Intuition, Alltagserfahrung, Erinnerung oder einer Mixtur aus dieser Art mehr oder weniger erfahrungsbasierter Introspektion eines an Kopräsenz interessierten Forschers oder einer entsprechend interessierten Forscherin beruhen und einer weitergehenden empirischen Grundlage entbehren.¹³ Näher an eine empirische Grundlegung führen dann schon Verfahren wie die teilnehmende oder intervenierende Beobachtung, mit denen sich Forschende auch lebenspraktisch und unter kaum ausschließbarem Handlungsdruck einer Kopräsenzsituation aussetzen (vom Lehnstuhl ins Feld) und die dabei gemachten Erfahrungen für die Forschung dokumentieren und aufbereiten. Aus ethnographischer Forschung dieser Art resultieren Aufzeichnungen unterschiedlichster Art, wozu längst nicht mehr nur Berichte, Tagebücher und Notizen gehören.¹⁴ Man könnte schließlich auch auf die Idee kommen, die an einer konkreten Kopräsenzsituation Beteiligten zu befragen, sei es durch Interviews, sei es durch Fragebogen, und auf diese Weise Daten generieren, die selbst Kommunikation dokumentieren (»Darstellungs-« vs. »Vollzugsmaterial«).¹⁵ Interaktionen im Fall von Interviews, lesbarkeitsbasierte Kommunikation im Fall von Fragebogen. Natürlich soll mit dieser Differenzierung von Zugängen nicht bestritten werden, dass nicht selten eine Kombination all dieser Zugangs- und Vorgehensweisen anzutreffen ist. Entsprechend sind in den neueren Methodendiskussionen in der Folge der Rezeption der Konversationsanalyse Methodenvielfalt und -ergänzung (insbesondere im Hinblick auf ethnographische Zugangsweisen) postuliert

worden. Audio- (und Video-)aufzeichnungen können gleichwohl nach wie vor als prinzipiell unverzichtbar gelten. Dieses Beharren auf Aufzeichnungen als einem Modus der registrierenden Datenkonservierung¹⁶ erscheint mit dem Postulat einer so wenig wie möglich durch Beobachtungsinteressen schon vorsortieren und -ausgewählten Datengrundlage forschungspraktisch gut motiviert. Methodologisch weiterführender ist aber der Effekt der Verfremdung, der mit der Erhebung von Daten in Form von Audio- (und Video-)aufzeichnungen verbunden ist und durch die Konfrontation mit solchen Daten provoziert wird. Tatsächlich entsteht schon mit der Audioaufnahme (und erst recht mit der Videoaufnahme) ein hoch künstliches Dokument, ein technisch voraussetzungsreiches Konstrukt, das die Vertrautheit mit der aufgezeichneten Situation auf wirksame Weise auflöst. Entgegen einer verbreiteten Unterstellung von »tape fetishism« oder Authentizitätsnostalgie« geht es nicht darum, eine Interaktionsepisode möglichst »naturgetreu« zu konservieren (wie es die Aufzeichnungssemantik in gewisser Weise nahelegt), sondern um ein »dekonstruktivistische[s] Manöver«, mit dem die fragliche soziale Wirklichkeit in einem neuen, so noch nicht wahrgenommenen Licht erscheint.¹⁷ Wer schon einmal die Aufnahme eines unspektakulären Alltagsgesprächs ohne eine irgendwie alltagsweltlich motivierte Neugier und ein irgendwie auf Bekanntheit beruhendes Interesse an den Beteiligten wieder und wieder abgehört hat, wird die Erfahrung machen, dass ihm oder ihr das Gehörte immer entstellter, fremder und verwunderlicher vorkommt. Diese Verfremdung muss ohne die Konfrontation mit Aufzeichnungen wahlweise durch theoretische Abstraktion (wie in der Systemtheorie), durch ein bewusstes Sich-dumm-Stellen gegenüber eingefahrenen Handlungsroutinen (wie in der Ethnomethodologie) oder durch eine bereits genügend unverständlich-»fremde« Gegenstandswelt (wie in der klassischen Ethnologie und -graphie) auf anspruchsvolle Weise erarbeitet werden. Auch das mehrfache Abhören von Interaktionsaufzeichnungen führt natürlich nicht wie von selbst zur Entdeckung der Probleme der Interaktion. Aber es provoziert gerade in seiner Künstlichkeit eine einzigartige Weise der Methodisierung. Der nächste elementare Schritt dazu ist die Transkription. Die Erfolgsgeschichte speziell der kalifornischen Konversationsanalyse ist eine Geschichte der Entdeckung von »Gesprächsdaten« in Form der Audioaufnahme und der nachträglichen Transkription des Gesprochen-Gehörten.¹⁸

III.

Landläufig versteht man unter dem Transkribieren einen Vorgang, bei dem (typischerweise sprachliche) Zeichen von einer (oftmals schriftlichen) Erscheinungsform in eine andere schriftliche Erscheinungsform übertragen werden. Das Ziel des Transkribierens ist also eine Umschrift, die Transkription. So kann z.B. ein Text für uns lesbar (gemacht) werden, den wir ansonsten nicht entziffern könnten, sei es, weil er in einer uns nicht vertrauten Schrift oder in einer unleserlichen Handschrift verfasst worden ist. Es kann aber z.B. auch ein Text für uns hörbar und aussprechbar gemacht werden, dessen lautliche Gestalt für uns verloren oder aus anderen Gründen noch nicht oder nicht mehr zugänglich ist. Das aus dem Fremdsprachenunterricht bekannte Internationale Phonetische Alphabet (IPA) ist eine Einrichtung, die solche Umschriften für alle Sprachen ermöglicht. Nichts davon trifft auf das in der Konversationsanalyse übliche Transkribieren gesprochener Sprache(n) in Form einer literalen Umschrift zu. Weder wird dadurch etwas verfügbar gemacht, was wir vorher nicht verstanden haben, noch wird dabei etwas umgeschrieben, was schon vorher geschrieben stand. Das in der Konversationsanalyse praktizierte und inzwischen mehr und mehr standardisierte Transkribieren gesprochener Sprache ist ein ganz und gar unpraktisches und lebenspraktisch entsprechend überflüssiges Unternehmen, bei dem auf den ersten Blick befremdlich anmutende Hybride zwischen Mündlichkeit und Schriftlichkeit entstehen. Mit genuin (druck-)schriftlichen Mitteln wird dabei versucht, Facetten von Mündlichkeit zu erhalten und einen Eindruck von Mündlichkeit beim Lesen hervorzurufen bzw. zu suggerieren und zu inszenieren. Man kennt solche Versuche auch aus der Literatur, so etwa aus dem Drama z.B. mit der Figurenrede und entsprechenden Hinweisen auf die Sprechsituation (»(Sie pflückt eine Sternblume und zupft die Blätter ab, eins nach dem andern)«) und vielleicht auch auf die Sprechweise (»Margarete (*halblaut*). Er liebt mich – liebt mich nicht«), aus der Fingierung gesprochener Sprache in Erzählungen (»Nun«, sagte ich, »wenn ich ein Taugenichts bin, so ist's gut, so will ich in die Welt gehen und mein Glück machen.«) und auch aus der sogenannten Mundartdichtung (»med ana schwaozzn dintn«). Im Gegensatz dazu orientiert sich die konversationsanalytische Transkription nicht an Phantasie und Erinnerung, sondern an der Ton- oder Videoaufnahme des Gesprochen-Gehörten. Ohne auf eine phonetische Umschrift zurückzugreifen, soll das Gesprochen-Gehörte möglichst so wiedergegeben werden, dass der mündliche Charakter des Sprechens-und-Zuhörens erhalten bleibt, genauer

gesagt: beim Lesen der Transkription unübersehbar wird. Dafür müssen die schriftorientierten Regeln der Interpunktion und Orthographie außer Kraft gesetzt und eine Reihe von Konventionen muss eingeführt werden, mit denen insbesondere Eigentümlichkeiten des Sprechens wie die Sprechgeschwindigkeit, die Betonung und weitere Besonderheiten des Tons und der Stimme, Abbrüche und Neustarts beim Sprechen, aber z.B. auch das gleichzeitige Sprechen mehrerer Sprecher:innen mit und durch Schrift dargestellt werden können. Die Praxis des Transkribierens kann heute auf eine jahrzehntelange Phase der Entwicklung und Verfeinerung zurückblicken, in die bereits Beobachtungen und Ergebnisse konversationsanalytischer Forschung auf vielfache Weise eingeflossen sind.¹⁹ Der Anfangsphase des Transkribierens, die in der Forschungssozialisation individuell nachvollzogen werden kann, haftet dagegen noch der Reiz an, mit dem Transkribieren den mit der Datenerhebung gegangenen ersten Schritt zur Verfremdung in der Datenaufbereitung gewissermaßen noch einmal zu forcieren. Das Transkribieren besteht ja gerade in dem Zwang zum detailgenauen Hinhören, der zum zigmalig wiederholten Abhören der fraglichen Aufnahmesequenz führt. Das fängt schon damit an, den Wortlaut möglichst authentisch zu erfassen, und führt wie von selbst zur Entdeckung immer weiterer Details des Sprechens-und-Zuhörens. Es ist in gewisser Weise das, was beim schnellen (Motiv-)Verstehen der Beteiligten zwar gehört und auch ausgewertet, aber in der Regel eben gar nicht bewusst wahrgenommen wird. Das Transkribieren bürstet dieses schnelle (Motiv-)Verstehen gegen den Strich²⁰ und ist damit, wie oft betont worden ist, der erste Schritt auf dem Weg zur Entdeckung eines relevanten Phänomens. In diesem Sinn wirft das Transkribieren viel mehr Fragen auf, als es beantwortet. Es versorgt die, die transkribieren, mit einer Überfülle an neuartigen Eindrücken von dem, was in der scheinbar so alltäglich-selbstverständlichen Welt der verbalen Interaktion fragwürdig und erklärenswert sein könnte.

In ihrer Einleitung in die Herausgabe ausgewählter Aufsätze von Gail Jefferson, die zur konversationsanalytischen Gründergeneration der frühen Jahre gehört, betonen Jörg Bergmann und Paul Drew die enorme Bedeutung des entdeckenden Transkribierens in anschaulichen Worten:

The process of transcribing requires the closest – one might say, intimate – acquaintance with the very intricate, complex, and often »messy« detail of talk-in-interaction. Transcription requires the closest attention to the most minute details of talk, in the first instance for their own sake. The transcriber attempts to capture, in written form, the details of precisely how

things are said, without yet knowing what will emerge from these details, what the details will amount to, or in what way the details will turn out to be significant or will matter for the analysis of those data. In listening for and transcribing the extraordinary complexity of what is often referred to as ordinary social talk, Jefferson was alerted to a great range of what she came to regard as the most significant interactional phenomena and practices.²¹

Wenn man diese Beschreibung ernst nimmt und ahnt, wie das Transkribieren die ersten konversationsanalytischen Entdeckungen animiert und beflügelt hat, ist es nicht verwunderlich, dass Gail Jefferson auch und gerade mit ihren Transkriptionsarbeiten maßgeblich an der Konstitution des »turn taking« als eines der zentralen Gegenstände konversationsanalytischer Forschung beteiligt war. Außerdem sieht man die Arbeit des Transkribierens dann nicht als lästige Pflichtübung der »Datenaufbereitung«, sondern als einzigartigen Schritt hin zur Explikation relevanter Analyseperspektiven. Wir sehen darin einen starken Auslöser für die Einsicht in die Unwahrscheinlichkeit der Interaktion.

Dabei darf man nicht vergessen, dass die Konversationsanalytiker:innen der ersten Stunde nicht der Linguistik angehörten (sondern der Soziologie) und dass sie kein spezielles Interesse an Sprache hatten (sondern an Alltagsinteraktion). Das sieht man heute noch am Prinzip der »Lesbarkeit des Transkripts«. Damit ist gemeint, dass die »Transkription [...] auch für Nicht-Linguisten lesbar sein [soll]«. ²² Auch wenn die Notationskonventionen mit den Jahren und auf der Grundlage der Beobachtungen und Ergebnisse immer differenzierter, formbezogener und linguistisch informierter geworden sind, haftet dem gesprächsanalytischen Transkribieren bis heute das Abenteuer an, eine in vielen Jahrhunderten für die Bedürfnisse der Kommunikation mit und durch Texte geschaffene Schriftsprache auf das gesprochene Wort im Gespräch anzuwenden. In das Medium der Schriftsprache übertragen, kehrt das gesprochene Wort dann gleichsam entstellt und entfremdet zurück. Alle, die ihre eigene Rede und Sprechweise in Alltagsgesprächen schon einmal transkribiert vor sich gesehen haben und dabei auf Pausen, auf »ähs« und »hms«, auf Satzabbrüche, Wiederholungen und Stockungen gestoßen sind, können davon ein Lied singen. Das Transkribieren leistet damit eine Stillstellung und zur Zurschaustellung des Gesprochenen, die die Aufmerksamkeit auf Probleme der Interaktion lenkt. Genau dazu bedarf es der Transkription und des mit ihr einhergehenden Verfremdungsaspektes, der sich weitgehend der Schrift verdankt und fast wie von selbst die Aufmerksamkeit auf die

Wunder und Wunderlichkeiten des Sprechens-und-Zuhörens lenkt. Erst in der durch und mit Schrift(-lichkeit) erreichten Dokumentation wird sichtbar, wie wenig selbstverständlich die gesprochene Alltagsinteraktion ist. Die Konversationsanalyse beutet diesen Effekt für analytische Zwecke und Anforderungen der von ihr kultivierten Art und Weise der Beobachtung aus. Sie entdeckt darin einen genuinen Forschungsgegenstand und hat ihre Transkriptions- und Notationskonventionen immer mehr und immer stärker den damit einhergehenden Rekonstruktionsanforderungen angepasst, was sich im Prinzip der »Relevanz« manifestiert, demzufolge beim Transkribieren solche Phänomene in den Mittelpunkt der Aufmerksamkeit rücken, die sich in der Entdeckung und Erforschung des Gegenstands der sozialen Interaktion als bedeutsam erwiesen haben. Entsprechend früh wurden mit dem Interesse am Sprecherwechsel Notationskonventionen für das überlappende Sprechen entwickelt und verfeinert, während erst in den letzten Jahren mit der Einbeziehung der Multimodalität der Interaktion Aspekte des visuell Zugänglichen auf der Grundlage von Videoaufzeichnungen notiert werden. Ein ähnlicher Beleg ist der Übergang zum Prinzip einer »formbezogenen Parametrisierung« interessierender Phänomene, dem folgend Sprechweisen nicht (mehr) mit (halb-)interpretativen Kommentaren (wie z.B. »empört«) wiedergegeben werden, sondern mit einer Identifizierung derjenigen prosodischen Phänomene, die eine solche Interpretation auslösen. Die Liste dieser (Weiter-)Entwicklungen des gesprächsanalytischen Transkribierens ließe sich vermehren. Es sollte aber auch so klar geworden sein, wie Forschung und Transkription hier ineinandergreifen und wie dabei die einfache Lesbarkeit der frühen Transkriptionen z.T. erheblich eingeschränkt wird, so dass im Endeffekt ein Text entsteht, der sich aufgrund seiner speziellen Notationskonventionen dem schnellen und routinisierten Lesen verweigert.

Es haftet der Idee des Transkribierens nach wie vor an (und steckt schon in der Rede von der Umschrift), dass die Transkription einen bloß sekundären, nachträglichen Charakter hat. Das ist nicht ganz falsch, wenn man daran denkt, dass es die Transkription des Gesprochen-Gehörten nicht geben würde, wenn es nicht irgendwann eine Interaktionsepisode gegeben hätte, in der die fraglichen Äußerungen aufgetreten wären, und nicht bereits zuvor jemand auf die Idee gekommen wäre, das fragliche Gespräch aufzuzeichnen. Die Transkription ist also, was ihren Status als Datum betrifft, ein echtes Sekundärdokument. Gleichwohl hat diese Vorstellung des bloß Sekundär-Nachträglichen der Transkription etwas Irreführendes, weil mit der Transkription etwas Neues in die Welt kommt, was es vorher nicht gegeben hat.

Man weiß inzwischen in der Phonetik, aber auch in der Linguistik der Alphabetschrift (und des Schriftspracherwerbs) sehr gut darüber Bescheid, wie sehr mit der (Alphabet-)Schrift die auditive Wahrnehmung von Sprachschall nicht nur geschult, sondern überhaupt erst kreiert wird. Die Transkription gesprochener Sprache hat deswegen nicht nur konservierenden, sondern eben auch produktiv-schöpferischen Charakter, was im Übrigen nicht nur für das linguistische Transkribieren oder die Anwendung der IPA, sondern auch schon für Darstellungen von (Sprach-)Schallwellen in Form von Oszillogrammen und Spektrogrammen gilt, mithilfe derer man z.B. gerne die physiologischen und akustischen Grundlagen der Phonetik illustriert.²³ Es geht in allen diesen Fällen zunächst um nicht mehr und nicht weniger als darum, die Welt des (Mit-)Gehörten auf eine neue Weise *sehen und lesen* zu lernen. Je erfolgreicher solche Modelle sind, desto mehr wirken sie dann auf die auditive Wahrnehmung des Gehörten selbst zurück: Wer die entsprechende Zeile in der Transkription gelesen hat, wird kaum umhinkönnen, die transkribierte Äußerung auch zu »hören«, wenn der Ton abgespielt wird (und die Transkription korrekt ist). Die Alphabetschrift, die vielleicht das erfolgreichste Modell dieser Art ist, lebt nicht zuletzt davon, dass sie suggeriert, dass mit ihr (bloß) aufgeschrieben wird, was zuvor gesprochen wurde – ein Mythos der Alphabetschrift, der inzwischen viel kommentiert worden ist²⁴ und von dem auch das gesprächslinguistische Transkribieren lebt. Anders als das literarische Fingieren von Mündlichkeit beruht das gesprächsanalytische Fingieren von Mündlichkeit allerdings nicht auf der Erinnerung an Gehörtes, nicht auf einer besonderen Vorstellungskraft von Oralität und auch nicht auf poetischem Gestaltungs- und Stilisierungsvermögen, sondern auf dem prinzipiell unendlich wiederholbaren Abhören der Tonaufzeichnung mittels eines Aufnahmegerätes (Rekorders). Die Audioaufnahme ist es, an der sich die gesprächsanalytische Transkription bewähren muss. Von dieser Orientierung am nicht nur einmal, sondern vielfach, nicht nur von einer Hörerin oder einem Hörer, sondern von vielen Hörerinnen und/oder Hörern Gehörten leben die Prinzipien des gesprächslinguistischen Transkribierens. Es wäre verheerend für die Konversationsanalyse und ihre Daten, wenn sie von diesem methodischen Credo der Transkription als eines Sekundärdokumentes abrücken wollte. Es wäre aber auch naiv, wenn man übersehen wollte, dass und wie das Lesbarmachen des Gesprächs mit der Transkription ein künstliches Konstrukt schafft, das die aufgezeichnete Gesprächssituation in vielfacher Hinsicht verzerrt – und auf genau diese Weise dazu beigetragen hat, einen bis dato unbemerkten Gegenstand zu entdecken und zu konstituieren.

IV.

Im Rückblick zeigt sich deutlich, dass und wie eine neue Technologie der Datenerhebung (der Audiorekorder, mit dem sich Telefongespräche aufzeichnen und praktisch unbegrenzt wiedergeben lassen) und eine darauf beruhende Technologie der Datenaufbereitung (die Transkription der Details des Gesprochenen, die auf der Möglichkeit der zigfachen Wiedergabe beruht) dazu beigetragen haben, einen Gegenstandsbereich zu entdecken, den es zuvor so nicht gab. In den 60er Jahren des 20. Jahrhunderts bedeutete es tatsächlich etwas Neues und Ungewohntes, sich mit Tonbandaufzeichnungen von (Telefon-)Gesprächen zu beschäftigen, sich diese Aufzeichnungen immer wieder anzuhören und dann damit zu beginnen, das vielfach Gehörte in einer bis dahin nicht bekannten Form möglichst nah am Gehörten wie oben beschrieben zu verschriftlichen (transkribieren). Man benutzte also die Technologie des Audiorekorders, um damit einen neuen Typus von Sekundärdaten zu gewinnen, der sich schnell gegenüber dem Primärdokument der Aufzeichnung verselbständigte. Es erinnert das an die bei Thomas Kuhn im Rückblick auf die Wissenschaftsgeschichte der Naturwissenschaften beschriebene Konstitution eines neuen Paradigmas. Im Fall der Konversationsanalyse war die neue Praxis der Datenherstellung, das Transkribieren, offenbar »neuartig genug [...], um eine beständige Gruppe von Anhängern anzuziehen, die ihre Wissenschaft bisher auf andere Art betrieben hatten«. ²⁵

Diejenigen, die auf diese Weise angefangen hatten, Gespräche mitzuschneiden und zu transkribieren, fanden diese neuartige Praxis offenbar aufregend genug, um sich darauf vertiefend einzulassen. ²⁶ Dazu kam das, was nach Thomas Kuhn die »Offenheit« eines neuen Paradigmas ausmacht: Mit den Transkriptionen wurde nicht nur viel mehr möglich an neuen Beobachtungen. Zugleich wurde viel mehr fraglich und problematisch, als man auf Anhieb überschauen konnte. Der Umgang mit Transkriptionen war deshalb offen genug, »um der neuen Gruppe von Fachleuten alle möglichen ungelösten Probleme zu stellen«. ²⁷ Transkriptionen erwiesen sich als Auslöser für die Einsicht in die mit dem Sprechen-und-Zuhören bearbeiteten Anforderungen. Alles das, was Gesprächsteilnehmer:innen im Verlauf einer Interaktion tun, ohne es überhaupt zur Kenntnis zu nehmen, wurde damit fraglich und erklärungsbedürftig im Sinne des Aufbrechens der unhinterfragten Gewohnheiten der Interaktionsroutine. Dazu zählt schon die Frage, wie Gesprächspartner:innen es überhaupt fertigbringen, dass sie nicht ständig zur gleichen Zeit reden, sondern sich tatsächlich über weite Strecken sehr systematisch

beim Reden und Zuhören abwechseln. Die, die Gefallen daran gefunden hatten, Gespräche aufzuzeichnen und zu transkribieren, und anfangen, sich als Konversationsanalytiker:innen zu verstehen und zu bezeichnen, entdeckten auf einen Schlag alle möglichen »Probleme«, über die wir im Gesprächsalltag normalerweise nicht nachdenken, die wir aber gleichwohl alltagspraktisch lösen. Dafür ist das Problem der Gesprächsbeendigung ein prominentes Beispiel. Emanuel Schegloff und Harvey Sacks haben es in einer ihrer ersten bahnbrechenden Arbeiten auf der Grundlage erster Einsichten in den Sprecherwechsel entdeckt.²⁸ Wenn es, vereinfachend gesagt, stimmt, dass *nach* dem Sprecherwechsel *vor* dem Sprecherwechsel ist, stellt sich analytisch das Problem, wie Gesprächsbeteiligte aus diesem rekursiven Prozess des Sichabwechselns überhaupt wieder herauskommen können, ohne willkürlich das Gespräch abubrechen. Es muss, anders gesagt, eine Vorrichtung geben, die es erlaubt, dass dieser Sprecherwechselmechanismus auf eine geregelte Weise ausgesetzt werden kann, ohne dass ein nächster »turn« erwartet werden kann bzw. werden muss. Dafür gibt es die Einrichtung des aus einer Paarsequenz bestehenden »terminal exchange«, der seinerseits in einem nicht unaufwendigen Prozess der Eröffnung der Beendigung vorbereitet werden muss (der dem besagten Artikel den Namen gegeben hat: »Opening up Closings«). Ohne hier weiter in die Details des Beendigens von Gesprächen zu gehen, ist es aufschlussreich, wie Schegloff und Sacks die Art von »(Beendigungs-)Problem« eingeschätzt haben, mit der sie es hier (erstmal) zu tun hatten. In einer eher beiläufig-ergänzenden (und um einen eingeklammerten Zusatz ergänzten) Randbemerkung arbeiten sie sich daran ab, dass sie ein »Problem« entdeckt haben, das den Beteiligten in der Regel nicht bewusst ist und das sie als solches gar nicht zur Kenntnis nehmen oder sonst wie reflektieren, das aber gleichwohl kein künstliches Beobachterkonstrukt ist, sondern sehr wohl ein »Problem« der Teilnehmenden:

In the ensuing discussion [...] it should be clearly understood that the ›closing problem‹ we are discussing is proposed as a problem for conversationalists; we are not interested in it as a problem for analysts except insofar as, and in the ways, it is a problem for participants. (By ›problem‹ we do not intend puzzle, in the sense that participants need to ponder the matter of how to close a conversation. We mean that closings are to be seen as achievements, as solutions to certain problems of conversational organization. While, for many people, closing a conversation may be a practical problem in the sense

that they find it difficult to get out of a conversation they are in, that problem is different from the problem of closing that we are concerned with. [...])²⁹

Das Zitat belegt, dass es um eine bestimmte Perspektive auf das Gesprächsgeschehen geht (»problems of conversational organization«), aus der heraus (auf einmal) explikations- und erklärungsbedürftig wird, was den Teilnehmenden gar nicht auffällt als problematisch, wenn sie nicht gerade in praktischer Hinsicht damit zu tun haben (weil sie z.B. ein praktisches Problem damit haben, aus einem konkreten Gespräch herauszukommen). Beendigungen sind deshalb Leistungen im Hinblick auf eine Aufgabe und Lösungen für ein Problem (»achievements«, »solutions to [a] certain proble[m]«), das der konversationsanalytische Beobachter bzw. die entsprechende Beobachterin entdeckt hat als ein typisches Problem konversationeller Organisation – auch wenn es den Beteiligten als solches nicht bewusst ist. Dass die Gesprächsbeendigung nicht nur in speziellen Fällen, sondern *immer* ein Problem ist, muss man tatsächlich erst zu sehen lernen.

In diesem Sinn geht es über das Beendigungsproblem hinaus in der Konversationsanalyse um alle möglichen Probleme, die damit anfangen, dass Gespräche nicht nur beendet, sondern auch eröffnet und aufrechterhalten werden müssen, und z.B. damit zu tun haben, dass wir uns nicht nur abwechseln, sondern dabei zugleich unsere Beiträge auch so gestalten, dass sie sich als Beiträge zu Gesprächsthemen verstehen lassen. Die Konversationsanalyse hat diese Probleme nicht interaktionstheoretisch abgeleitet und vorgegeben, sondern in der Auseinandersetzung mit den Daten (»from the data themselves«)³⁰ – und d.h. im Erstellen und im Arbeiten mit den Transkriptionen – entdeckend rekonstruiert. In diesem Ineinandergreifen von Empirie und Theorie besteht die konversationsanalytische Innovation und der Grund der außergewöhnlichen Erfolgsgeschichte dieses Ansatzes. Der konversationsanalytische Problembegriff bringt das auf empirisch gehaltvolle Begriffe, was Luhmann die Unwahrscheinlichkeit der Kommunikation nennt und worauf wir jetzt auch theoretisch noch näher eingehen wollen.

V.

In einem seiner früheren Aufsätze, dessen Grundgedanken auch in späteren Arbeiten in fast unveränderter Form wieder auftauchen, verweist Luhmann auf eine allgemeine »theoretische Intention«, die er mit »einer These der Unwahrscheinlichkeit« verbindet.³¹ Was ist damit gemeint? Ausgangspunkt ist

die Unterscheidung zweier Theorieoptionen. Während die eine darin besteht, nach den »Möglichkeiten der Verbesserung der Verhältnisse« zu fragen, geht es bei der anderen (und favorisierten) Option »um eine theoretische Vorfrage aller Verbesserungen«, die sich darauf richtet, wie überhaupt die Verhältnisse, so wie sie sind, möglich und wahrscheinlich werden konnten. Sie, die relevante Option, »beginnt« deshalb »mit einer These der Unwahrscheinlichkeit« sozialer Ordnung, so dass zuallererst die Frage auftaucht, »wie Zusammenhänge, die an sich unwahrscheinlich sind, dennoch möglich, ja hochgradig sicher erwartbar werden«. Es geht in genau diesem Sinne darum, »die Routineerwartungen und die Sicherheiten des täglichen Lebens« weitgehend aufzulösen, wozu es »einer sozusagen contra-phänomenologischen Anstrengung« bedarf. Nur dann lässt sich behaupten, dass Kommunikation unwahrscheinlich ist, »obwohl wir sie jeden Tag erleben, praktizieren und ohne sie nicht leben würden.« Ein wichtiger Schritt hin zum Begreifen dieser »unsichtbar gewordenen Unwahrscheinlichkeit« besteht darin, »Kommunikation nicht als Phänomen, sondern als Problem« aufzufassen.³²

Luhmann führt das in dem vorliegenden Aufsatz auf allgemeine Weise am Beispiel dreier Grundprobleme der Kommunikation vor Augen, die so gewählt sind, dass sie auf drei »Einrichtungen« verweisen, die es möglich machen, dass diese Grundprobleme überwunden werden können. Die drei Grundprobleme begründen jeweils einen eigenen Typ von Unwahrscheinlichkeiten, die sich wechselseitig verstärken: eine Unwahrscheinlichkeit des *Verstehens* (wieso sollte überhaupt einer verstehen, was der andere meint), eine Unwahrscheinlichkeit des *Erreichens* Anderer, wenn Kommunikation den Kreis Anwesender überwinden können soll, und eine Unwahrscheinlichkeit des Erfolgs der Kommunikation im Sinne des Annehmens und des *Erfolgs* der damit verbundenen Setzungen. Auf diese drei Probleme reagiert, so die allgemeine These, die Entwicklung der »Medien«: mit *Sprache* auf das Problem des Verstehens, mit *Verbreitungsmedien* auf Probleme der Erreichbarkeit und mit *symbolisch generalisierten Kommunikationsmedien* auf das Problem des Erfolgs. Medien können auf diese Weise erklären helfen, wie die Unwahrscheinlichkeiten der Kommunikation durch und mit Kommunikation immer wieder überwunden und in diesem Sinne geheilt werden können, anstatt als Schwellen der Entmutigung Kommunikation gar nicht erst zustande kommen zu lassen. Damit sind die theoretischen Grundlinien einer allgemeinen Theorie der Medien- und Gesellschaftsdifferenzierung bestimmt, die »ein breit ausgezogenes Forschungsprogramm« hervortreten lassen. So elegant die Systematik von Problemexplikation und Mediendifferenzierung entwi-

ckelt wird, so deduktiv und abstrakt wirkt freilich auch die Umsetzung der Unwahrscheinlichkeitsthese.

Im Gegensatz dazu tritt die hier in den Mittelpunkt gerückte Unwahrscheinlichkeit der Interaktion auf eine höchst anschaulich-konkrete Weise hervor. Das hat damit zu tun, dass sie sich nicht auf das Gesamte dessen bezieht, was Kommunikation und Gesellschaft ausmachen, sondern nur auf den eingeschränkten Bereich der Kommunikation unter Anwesenden, d.h. der kopräsenzbasierten Interaktion. Und vor allem hat es damit zu tun, dass die Unwahrscheinlichkeit der Interaktion nicht einer abstrakten Theorieoption folgt, sondern einer methodisierten Fremdheitserfahrung im Umgang mit aufwendig erhobenen und aufbereiteten Daten. Die »contra-phänomenologisch[e] Anstrengung«, von der Luhmann spricht (s.o.), ist im Fall der Konversationsanalyse empirisch motiviert im Umgang mit einem auf spezifische Weise erhobenen und aufbereiteten Datenmaterial. Vor allem die Transkription hat sich im Nachhinein als ein Auslöser für die Explikation von Problemen erwiesen. Beobachtete und transkribierte Erscheinungsformen von Interaktion (und sei es nur ein Räuspern, oder besser gesagt: *gerade* ein Räuspern) verlangen danach, entdeckend zu erklären, auf welches Problem sie reagieren. Hermeneutisch formuliert ist es der Konversationsanalyse gelungen, die Fragen zu finden und zu formulieren, auf die die Transkriptionen in ihren Details immer schon die Antworten liefern. In diesem Sinne kommen Probleme in großer Zahl an die Oberfläche, sobald die Transkriptionsmaschine einmal in Gang gekommen ist. Mit der kalifornischen Folkrockgruppe *We Five*, deren Sängerin Beverly Bivens in einer Videoaufnahme ihres Songs *You Were On My Mind* voller Freude und Ausgelassenheit singt: »I got troubles, whoa-oh, I got worries, whoa-oh«,³³ mag man sich die Konversationsanalytiker:innen vorstellen, wie sie im Angesicht ihrer Transkriptionen jubeln: »I got problems, whoa-oh, I got problems, whoa-oh«. Vor dem Hintergrund von mehreren Jahrzehnten konversationsanalytischer Forschung ist es heute vergleichsweise einfach, diese Probleme zu explizieren und zu systematisieren als Probleme des Anfangens und Aufhörens (Wann geht's los? Wann hört's auf?), des Sichabwechselns (Wer kommt als Nächste oder Nächster?), der Themenwahl und -bearbeitung (Was kommt als Nächstes?), der Selbst- und Fremddarstellung (Wer sind wir?), der Situierung (Wo sind wir?) und der Rahmung (Was geht hier vor?). Der Prozess der Problemexplikation ist damit nicht schon vollumfänglich abgeschlossen, zumal man sehen muss, dass sich hinter diesen »big six« der Konversationsanalyse unzählige Unter- und Unterunterprobleme verstecken, sobald man einmal anfängt, sich näher mit ihnen

zu beschäftigen (worauf wir in diesem Kapitel noch zurückkommen). Im vorstehenden Zusammenhang kommt es darauf an, dass vor dem Hintergrund in die Einsicht dieser Probleme die Alltagsroutine und Selbstverständlichkeit der Interaktion zugunsten des Axioms der Unwahrscheinlichkeit der Interaktion in verschiedenen Richtungen aufgelöst werden kann: kopräsenzbasierter Interaktion erscheint als höchst problematische, stets vom Scheitern bedrohte, überaus riskante, mit einem Wort: höchst unwahrscheinliche Kommunikation, die vor Problemen nur so wimmelt. Es dürfte kein Zufall sein, dass diese Einsicht in die Unwahrscheinlichkeit der Interaktion auf der Grundlage verbaler Daten zustande gekommen ist. Die Konfiguration von Kopräsenz durch Sprache darf man sich als einen wesentlichen Problemlösungsschub für Interaktion vorstellen, mit dem Kopräsenz auf eine Weise möglich und wahrscheinlich geworden ist, wie es vor der Errungenschaft von Sprache undenkbar gewesen sein dürfte. Die Aufgaben, die wir dazu noch skizzieren werden, machen deshalb auf Probleme aufmerksam, die dazu beigetragen haben, Sprache so zu entwickeln und zu formen, dass sie mit den neuen Herausforderungen einer immer weitergehenden sprachlichen Konfiguration von Kopräsenz zurechtkommen und Schritt halten konnte. Die Grammatik und das Lexikon natürlicher Sprachen sind der Ausdruck dieser Formentwicklung.

Wenn man allerdings an phylo- und ontogenetisch frühe Formen von Kopräsenz herankommen will, muss man die fraglichen Probleme der Interaktion so formulieren, dass damit der Anspruch erhoben werden kann, hinter das Auftauchen von Sprache und die Konfiguration von Kopräsenz durch Sprache zumindest perspektivisch zurückzutreten. Man sieht dann sofort, dass es Problemstellungen gibt, die eng an die Entwicklung und den Einsatz von Sprache geknüpft sind. Hierhin gehören offenkundig der Sprecherwechsel und die Themenorganisation, also die Aufgaben des Sichabwechselns und des Beitragens zu einem Thema. Und man sieht auch in den anderen Fällen, dass sich Gestalt und Charakteristik der Aufgabenstellung verändern, sobald – und in dem Maße, in dem – Sprache bei der Erledigung dieser Aufgaben im Spiel ist. Schließlich kommt hinzu, dass unser Modell der »big six« empirisch nur für den rezenten Bereich der verbalen Interaktion gedeckt ist. Wenn wir Probleme formulieren wollen, die auch mit Blick auf phylogenetisch frühe Formen von Kopräsenz plausibel sein sollen, müssen wir zwangsläufig über den engen Forschungsstand der Konversationsanalyse und Gesprächsforschung hinausgehen. Wir wollen dazu mit der Besprechung einer Aufgabe beginnen, die in der einschlägigen Forschung lange vernachlässigt worden ist, weil sie die

vielleicht am wenigsten durch Sprache geprägte Aufgabe darstellt. Das ist die Aufgabe der Situierung.

Das Problem der Situierung: Ko-Orientierung, Ko-Ordination und Ko-Operation

Das Problem der Situierung besteht in der Verankerung von Kopräsenz in der sinnlich wahrnehmbaren Umgebung der Anwesenden.³⁴ Dieses Problem hat, wie wir in diesem Abschnitt erläutern werden, sehr viel mit Raum und Räumlichkeit zu tun und ist als eigenständiges Problem in der Konversationsanalyse und Gesprächsforschung lange nicht gesehen worden.³⁵ Es ist der »Interaktionsraum«, um dessen Herstellung es bei der Situierung geht. Die Situierung hat bis heute im Schatten anderer Probleme gestanden, die mit der Prozessualität der verbalen Interaktion zu tun haben. Dass Kopräsenz mit der »Sequentialität« der Interaktion eine eigene soziale Zeitlichkeit entwickelt hat, ist schon früh gesehen und empirisch wie theoretisch weiterentwickelt worden. Dass sie auch eine eigene Räumlichkeit hat, ist erst in den letzten rund fünfzehn Jahren unter dem Schlagwort des Interaktionsraums vermehrt erforscht worden. Wir werden unsere Bemerkungen zur Situierung deshalb mit einem kurzen Rückblick auf die Wiederentdeckung des (Interaktions-)Raumes einleiten, bevor wir darauf eingehen, dass die Situierung als Interaktionsproblem eine eigene grundlegende Unwahrscheinlichkeit der Interaktion begründet. Diese Unwahrscheinlichkeit reicht nicht zufällig an den Übergang von individueller Wahrnehmung zu sozialer Kopräsenz heran, insofern es zunächst um einen Raum für wechselseitig geteilte Aufmerksamkeit geht.

Eine nähere Betrachtung der Situierung führt darauf, dass der Interaktionsraum analytisch differenziert werden kann in einen Wahrnehmungs-, Bewegungs- und Handlungsraum. Die Aufgabe der Situierung umfasst entsprechend mit der Herstellung des Interaktionsraums die Schaffung von Wahrnehmungsräumen (Ko-Orientierung), Bewegungsräumen (Ko-Ordination) und Handlungsräumen (Ko-Operation). Auch wenn diese Räume empirisch in der Regel ineinander übergehen, ist es weiterführend, sie analytisch zu differenzieren. Wie wir zeigen wollen, dürfte im Übergang zur Ko-Operation der Schlüssel zur Entfaltung der uns bis heute geläufigen Interaktionsprobleme gelegen haben. Ko-Operation ist auch ohne Sprache (bis heute) möglich, aber es liegt auf der Hand, dass von hier aus weitere Anforderungen an Profil gewinnen, die mit der Markierung und Inszenierung der

Grenzen der Interaktionsepisode zu tun haben (Eröffnung und Beendigung), mit dem Sichabwechseln und dem Beitragen zum Bezug auf Welt (Sprecherwechsel und Themenorganisation) und mit der Selbstvergewisserung als Handlungskollektiv im kommunikativen Haushalt von Gemeinschaft und Gesellschaft (Selbst- und Fremddarstellung sowie Rahmung). Als sich allmählich profilierende Interaktionsprobleme dürften diese Folgeanforderungen der Ko-Operation die Schrittmacher der Evolution des Sprechens-und-Zuhörens sein.

I.

Wann immer Kopräsenz zustande kommt, impliziert das eine interaktive Verarbeitung sinnlicher Wahrnehmung. Das fängt mit der für Kopräsenz konstitutiven Reflexivität der Wahrnehmung der Beteiligten an: Mit dem Wahrnehmen des Wahrgenommenwerdens läuft die Interaktion an, indem sie individuelle Wahrnehmungsakte überführt in eine genuin soziale Situation wechselseitig geteilter Aufmerksamkeit. Damit ist offenkundig ein Prozess der Selektion verbunden, der vieles ausschließt, was die Beteiligten noch wahrnehmen und weiter wahrnehmen könnten, und manches einschließt, was die Beteiligten nur aufgrund dieser Beschränkung (und Konzentration) weiterverfolgen können. Ein wesentlicher Bestandteil der interaktiven Verarbeitung von Wahrnehmung besteht also in der Selektivität des Zugriffs, abstrakter gesagt: in der Reduktion der Komplexität des für die Beteiligten aufgrund ihrer Sinnesorgane Wahrnehmbaren. In jeder Interaktion muss entsprechend auf irgendeine Weise geklärt werden, was in einer konkreten Situation der Begegnung aus der Fülle des sinnlich Wahrnehmbaren (Sicht-, Hör-, Greif- und Tast-, Riech- und Schmeckbaren) auch tatsächlich für das (weitere) Zusammensein relevant werden soll – und was bis auf Weiteres außen vor bleiben kann. Wir nennen diese Aufgabe Situierung, weil sie dazu beiträgt, Kopräsenz mit den als relevant erachteten Aspekten der Situation zu verbinden und die Interaktion entsprechend situativ zu verankern. Das *Zeigen*, mit dem sich Gesprächsbeteiligte auf etwas in ihrer wahrnehmbaren Umwelt aufmerksam machen, ist vielleicht der auffälligste Ausdruck dieser Art von Situierung und situativer Verankerung. Zeigen hat unmittelbar mit Wahrnehmung zu tun; es ist das Medium wechselseitig geteilter Wahrnehmungen, in dem Aspekte der Umgebung überhaupt für die Interaktion relevant werden können. Zeigegesten und -ausdrücke (Deiktika) sorgen dafür, dass Wahrnehmungen selbst wahrnehmbar werden. Es hat deshalb den Anschein, als ob sie

»indexikalisch« auf die Situation zurückverweisen würden, in der sie geäußert werden. Tatsächlich tragen sie aber maßgeblich dazu bei, dass so etwas wie eine Situation überhaupt zustande kommt. Situierung meint deshalb streng genommen nicht die Verankerung in einer bereits zur Verfügung stehenden Situation (auf die sich zeigen ließe), sondern die Herstellung der zur Interaktion passenden Situation (durch Zeigen). Situative Verankerung schafft also erst die Art von situativem Hintergrund, der gerade interaktiv benötigt wird.

Diese Richtigstellung mag kleinlich erscheinen, vielleicht auch nach einer überflüssigen Verkomplizierung des Sachverhalts klingen. Sie ist aber notwendig, wenn man dem Mythos Kopräsenz entgehen will; zu diesem Mythos gehört die Annahme einer gegebenen Raum-Zeit-Stelle, auf die die Anwesenden im Sinne der »Sprechsituation« zurückgreifen können (Raum-Zeit-Apriori).³⁶ Entsprechend sind die Situation und mit ihr der Interaktionsraum oftmals eher vorausgesetzt als untersucht worden.³⁷ Das gilt konzeptionell mit wenigen Ausnahmen für die linguistische Pragmatik, wie man speziell daran studieren kann, wie anknüpfend an Karl Bühler der Raum- und Ortsbezug des Sprechens- und Zuhörens als Aspekt der Sprechsituation kanonisiert worden ist. Als vorausgesetztes »Hier« des Sprechenden bildet dieser Raum- und Ortsbezug bekanntlich eines der drei Fundamente der »Origo«,³⁸ die für das Verstehen sprachlicher Ausdrücke eine grundlegende Bedingung darstellt. Entdeckt und erforscht wurde diese Grundbedingung natürlicher Sprachen nicht zufällig am Beispiel der Deixis, und bis heute reißen die Versuche nicht ab, das von Bühler pionierhaft in die Sprachbetrachtung eingeführte »Zeigfeld der Sprache« theoretisch und empirisch zu fassen.³⁹ Dieser im deutschsprachigen Raum maßgeblich durch Bühlers Zeigfeldlehre etablierte Ausgangspunkt ist allerdings nicht frei von problematischen Vorannahmen. Dazu gehört namentlich der Ego-Zentrismus der Bühler'schen Konzeption,⁴⁰ mit deren Rezeption die Sicht auf die Sprechsituation oftmals kognitivistisch eingeengt worden ist.⁴¹ Damit hängt zusammen, dass die Sprechsituation selbst als eine gegebene Konstellation betrachtet wurde, die man dem Sprechen (und Zuhören) gleichsam gegenüberstellen kann,⁴² wie es nicht nur der Sprachfixierung der linguistischen Pragmatik geschuldet ist, sondern auch dem »correlational drive« der frühen Soziolinguistik entsprach.⁴³ Was ist daran (so) problematisch? Problematisch ist, dass es bei dem hier in Betracht kommenden Raum nicht um etwas geht, das sich gleichsam von außen als gegebenes Datum erfassen und hinzufügen lässt. Man kann sonst nicht sehen, dass die Leistung der Deixis gerade darin besteht, Wahrnehmung für die Interaktion überhaupt zugänglich zu machen und damit der Interaktion

zur Verfügung zu stellen. Die Annahme einer gegebenen »Sprechsituation« mit räumlichen Parametern verschleiert diesen Mechanismus der Wahrnehmungswahrnehmung, indem sie die Alltagsperspektive der Beteiligten einnimmt, für die Raum und Situation nicht konstruiert werden müssen, sondern bereits »da« sind, so dass man darauf ohne weiteres verweisen kann. Wenn man bei dieser Sichtweise stehen bleibt, kann man freilich nicht erklären, was am Zeigen so erstaunlich und so unwahrscheinlich ist.⁴⁴

Man kann die Problematik der Annahme einer gegebenen Sprechsituation nicht nur anhand der Entwicklung der linguistischen Deixiskonzeptionen illustrieren, sondern auch am Sprechen-und-Zuhören beim (herkömmlichen) Telefonieren. Auf den ersten Blick geht in diesem Fall der gemeinsame Sprech- und Zuhörerraum verloren. Physisch anwesend sind die Beteiligten beim Telefonieren ja gerade nicht. Nur wenn man Anwesenheit als eine im Medium der Wahrnehmungswahrnehmung hergestellte soziale Errungenschaft begreift, kann man sehen, dass Anwesenheit eben auch am Telefon tatsächlich hergestellt wird: Mit dem Abnehmen des Hörers nach dem Klingelzeichen können die Beteiligten wahrnehmen, dass sie sich wahrnehmen, auch wenn diese Wahrnehmungswahrnehmung ganz auf Hörbarkeit beschränkt ist. Für den Interaktionsraum der Beteiligten bedeutet das, dass er ganz und gar auf Hörbarkeit angewiesen ist, d.h. im Medium der auditiven Wahrnehmung entsteht und visuell und haptisch unzugänglich ist. So bleiben der Sprechort und z.B. Bewegungen der Beteiligten beim Telefonieren für die Interaktion am Telefon irrelevant, solange sie keine hörbaren Spuren hinterlassen (und z.B. thematisiert werden) und damit in den auditiv konstituierten Interaktionsraum hineinragen.⁴⁵ Man kann an diesem Spezialfall der Interaktion lernen, dass und wie der Interaktionsraum und mit ihm die Sprechsituation tatsächlich erst in und mit Interaktion geschaffen werden, so dass sie nicht als statisch, sondern als dynamisch, nicht als vorgegeben, sondern als emergent zu betrachten sind. Der Interaktionsraum hat in diesem Sinne eine eigentümliche Beschaffenheit: Er »ist« nicht, sondern »ereignet sich«.⁴⁶ Deshalb lässt er sich nicht durch eine Reihe vermeintlich gegebener Parameter von außen bestimmen. Als in und mit Interaktion jeweils fallspezifisch konstituierte Wirklichkeit muss er vielmehr als eine Hervorbringung empirisch rekonstruiert werden. Das Postulat eines eigenständigen Problems der Situierung dient dazu, diese Rekonstruktionsarbeit anzuleiten und zu systematisieren. Mit dem Postulat der Situierung wird der Raum vom Explanans zum Explanandum.

In der Interaktionsforschung ist der Raum lange vernachlässigt worden. Eine Theoretisierung der räumlichen Dimension von Kopräsenz, die mit der Theoretisierung ihrer zeitlichen Dimension mithalten könnte, steht nach wie vor noch aus. Man muss dazu nur daran denken, in welcher vielfältiger wie grundlegender Weise die Bedeutung von Zeit und Zeitlichkeit für die Interaktion in Anlehnung an das konversationsanalytische Konzept der Sequentialität elaboriert worden ist. Die Gründe für diese Vernachlässigung des Raumes dürften vor allem mit methodologischen und technologischen Aspekten der Datenkonstitution zu tun haben: Die Praxis der Transkription des Gesprochen-Gehörten hat maßgeblich dazu beigetragen, die sprachliche Interaktion aus ihrem multimodalen Entstehungskontext herauszulösen und als Gegenstand gegenüber ihrer situativen Einbettung zu verselbständigen. Mit den transkriptionsbasierten (und genau in dieser Hinsicht höchst innovativen) Untersuchungen der Konversationsanalyse⁴⁷ sind die nicht hör-, sondern nur sichtbaren Erscheinungsformen der Interaktion nach und nach aus dem Blick geraten. Sie waren zuvor in der Tradition der »context analysis« in ihrer Relevanz durchaus gesehen und auf der Basis von Videoaufzeichnungen auch bereits detailliert untersucht worden.⁴⁸ Mit dem Aufschwung und Erfolg der transkriptionsbasierten Konversationsanalyse ist diese Forschungstradition dann aber in den Hintergrund getreten. Der Raum gehörte damit wie vieles andere an »nonverbaler Kommunikation« zu dem, was man als situativen Hintergrund von Fall zu Fall ergänzen, aber nicht eigenständig analysieren konnte; in der Transkription tauchte er nicht auf, solange er nicht auch von den Beteiligten ausdrücklich thematisiert wurde.⁴⁹ Er blieb damit oftmals im wahrsten Sinne des Wortes »unsichtbar«, weil sich die Herstellung des Interaktionsraums in sehr vielen Fällen unterhalb der Schwelle der Verbalisierung abspielt. Diese Sichtweise gilt heute mit dem Aufschwung und der Verbreitung von Videoaufzeichnungen als Datenbasis und der Betonung der Multimodalität der Interaktion zu Recht als überholt. Es kommt hinzu, dass sich die Forschung inzwischen vermehrt auch solchen Settings zugewandt hat, in denen der Raum eine unübersehbar wichtige Rolle spielt und infolgedessen nicht ignoriert werden kann (worauf wir noch zurückkommen werden). In dem Maße, in dem die Videoaufzeichnung und ihre Sekundärdokumente (z.B. in Form von Standbildern) entsprechend neben die Transkription des Gesprochen-Gehörten als zentrales Bezugsdatum treten, tritt der Raum in Erscheinung und in seiner Gemachtheit für die Interaktion sichtbar und analysierbar hervor. Als Interaktionsraum (»interactional space«)

ist er eine interaktive Hervorbringung (»interactive achievement«), die es als solche zu untersuchen gilt.⁵⁰

II.

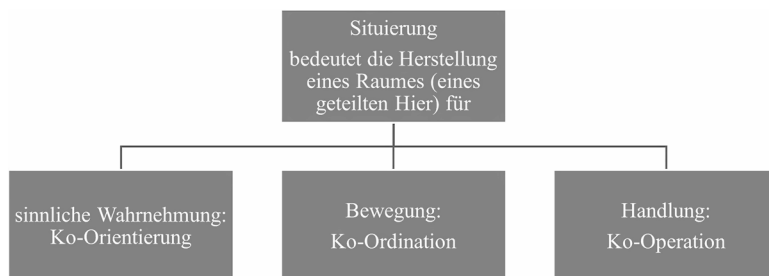
Der Interaktionsraum ist ein Element der Situation und als solches an die Wahrnehmungsleistungen der Anwesenden gebunden. Sobald die Interaktionsbeteiligten als sozial Anwesende auf irgendeine Art und Weise in Kopräsenz zusammenkommen, etablieren sie auf der Grundlage reflexiver Wahrnehmung einen gemeinsamen Aufmerksamkeitsfokus. Er geht in vielen Fällen bereits aus der Konfiguration der körperlich-räumlich eingenommenen Positionen hervor. Diese Konfigurationen tragen maßgeblich zur Konstitution des Interaktionsraums bei, ohne dass darüber in der Regel nachgedacht oder gar gesprochen werden muss. Der Interaktionsraum ist so selbstverständlich wie die Positionierungen, die dazu gehören, wenn man sich trifft, begegnet, zusammenkommt oder sonst wie einander in das Feld gegenseitiger Aufmerksamkeit gerät (»togetherness«). Die bekannteste und gewissermaßen kanonische Konfiguration dieser Art ist ein Arrangement, bei dem die Interaktionsteilnehmer:innen wechselseitig die Vorderseite ihrer Körper und damit ihre Wahrnehmungs- und Greiforgane aufeinander ausrichten und abstimmen (»face-to-face«).⁵¹ Auch wenn diese Konfiguration dem gesamten Gegenstandsbereich als »face-to-face interaction« den Namen gegeben hat, ist sie nicht alternativlos. Die Anwesenden können je nach Erfordernis der jeweiligen sozialen Praxis (darin einbegriffen: je nach Anzahl der Teilnehmer:innen) und im Prinzip jederzeit diese Konfiguration dynamisch verändern und dabei neue und andere Interaktionsräume konstituieren, die nicht mehr auf »Face-to-face«-Konstellationen, sondern beispielsweise auf »Side-by-side«-, »Face-to-back«- oder Kreis- und Halbkreis-konstellationen beruhen, um nur ein paar besonders prominente Positionierungen zu nennen. Es kommt hinzu, dass auch die »Face-to-face«-Konfiguration nicht zwangsläufig statisch ist. Kopräsenz umfasst nicht nur den Typus des »stationären« Aufenthalts der Beteiligten an einem und dem gleichen Ort in einer und der gleichen Konstellation (wie z.B. beim Gespräch am Tisch), sondern auch einen Typus, bei dem die Beteiligten »mobil« sind, d.h. sich fortbewegen und dauerhaft in Bewegung sind (wie z.B. beim gemeinsamen Spaziergehen oder einer Stadtführung). Zwischen diesen Eckpunkten fixierter und flottierender Anwesenheit gibt es vielfältige Abstufungen und Differenzierungen mit fließenden Übergängen. Man darf sich den Interaktionsraum deshalb

grundsätzlich nicht als etwas ein für alle Male Gegebenes vorstellen, sondern muss den Interaktionsraum unbeschadet der Materialität und Physis eines konkreten Raumes als eine fluide, jeweils von Augenblick zu Augenblick (lokal) im Medium geteilter Wahrnehmung emergente soziale Konstruktion vergegenwärtigen. Als in und mit Kopräsenz hervorgebrachte Konstruktion wird der Interaktionsraum *rekonstruierbar*. Genau das ist der Anspruch, den wir mit der Aufgabe der Situierung verbinden.

Die Situierung erfolgt in der Regel unterhalb der Verbalisierungs- und Bewusstheitsschwelle der Anwesenden, wozu Architektur und Mobiliar einen sehr großen Beitrag leisten (worauf wir noch zurückkommen). Es ist deshalb heuristisch weiterführend(er), sich das Problem der Situierung an Fällen zu vergegenwärtigen, in denen die Konstruktion des Interaktionsraums sichtbar(er) hervortritt und empirisch greifbar(er) wird. Das ist der Fall, wenn wir an den schon erwähnten Typus von Kopräsenz denken, bei dem die Beteiligten zusammen in Bewegung sind und die Herstellung des Interaktionsraums (= die Situierung) als Aufgabe im Vordergrund der Interaktion steht, so dass der sprachliche, sensomotorische und architektonische Situierungsaufwand in seiner Komplexität hervortritt. Solche mobilen Settings sind mit der Zuwendung zum Interaktionsraum in den letzten Jahren vermehrt untersucht worden.⁵² Was in Untersuchungen dieser Art anhand von Detailbeobachtungen von Interaktionsepisoden unterschiedlicher Art immer wieder hervortritt, ist die Komplexität des Situierungsproblems. Der Interaktionsraum ist komplex, und mit ihm ist es die Aufgabe der Situierung. Schon »face-to-face« ist deshalb nicht immer und einfach »face-to-face«, sondern eine Chiffre für eine prototypische Verkörperung wechselseitiger Wahrnehmbarkeit und ihrer Manifestation in Form von Blick(-kontakt), Mimik und Gesichtsausdruck, Nähe und Zuwendung, Körperhaltung und -ausrichtung, Gestik und Bewegung, kurz gesagt: ein Kürzel für eine multimodal im Zusammenspiel von Proxemik und Kinesik zustande gebrachte Konfiguration von Kopräsenz. Man sieht mit dem Aufschwung videobasierter Interaktionsforschung heute klarer, dass und wie Kopräsenz als Konfiguration anwesender Körper Raum zugleich impliziert und produziert: als *Wahrnehmungsraum*, der auf einem Minimum geteilter Aufmerksamkeit beruht, als *Bewegungsraum*, der nicht erst in den erwähnten mobilen Settings hervortritt, sondern grundsätzlich durch die Mobilität menschlicher Sensorik konstituiert wird, und als *Handlungsraum*, der dadurch zustande kommt, dass Wahrnehmungen und Bewegungen in der Regel eingebettet sind in eine bestimmte soziale Praxis und den mit ihr verbundenen Aktivitätstyp. Wahrnehmungs-, Bewegungs- und Handlungs-

raum machen zusammengekommen den Interaktionsraum aus. Für diese (Teil-)Räume gilt, was auch für den Interaktionsraum insgesamt gilt: Sie sind nicht einfach mit Kopräsenz gegeben, sondern werden in und mit Kopräsenz hervorgebracht. Um diese Hervorbringungsleistung zu betonen, sprechen wir von der Aufgabe der Situierung.⁵³ Sie besteht dann darin, unter Anwesenden geteilte Räume für Wahrnehmung, Bewegung und Handlung zu schaffen. Die Beteiligten müssen sich, anders gesagt, auf ein gemeinsames Hier für Wahrnehmung, Bewegung und Handlung verständigen. Um diese grundsätzlich triadische Struktur zu betonen, die sich an Vorarbeiten zu einer »Sozialpsychologie des Raumes und der Bewegung« anlehnen kann,⁵⁴ sprechen wir von drei Unteraufgaben, die mit der Situierung prinzipiell verbunden sind. Dabei handelt es sich um Ko-Orientierung, Ko-Ordination und Ko-Operation:

Abb. 13: Die Komplexität der Situierung (eigene Darstellung)



Wann immer Kopräsenz zustande kommt, müssen die Aufgaben der Ko-Orientierung, Ko-Ordination und Ko-Operation irgendwie bearbeitet werden, auch wenn die Erledigung dieser Teilaufgaben typischerweise im Hintergrund der Interaktion abläuft, ohne besondere Aufmerksamkeit auf sich zu lenken. Dabei handelt es sich um genuin interaktive Anforderungen, die sich auf zwischen mindestens zwei Anwesenden *wechselseitig* geteilte (reflexive) Aufmerksamkeit, auf zwischen mindestens zwei Anwesenden *wechselseitig* aufeinander abgestimmte Bewegungen und zwischen mindestens zwei Anwesenden *wechselseitig* aufeinander bezogene Handlungen (und »Handlungsspiele«) beziehen. Das Präfix »Ko-«, das in die Benennung dieser Unteraufgaben eingegangen ist, ist deshalb von grundlegender Bedeutung. Die Aufgaben richten sich streng genommen nicht an die beteiligten Individuen, die für sich genommen natürlich ständig mit Wahrnehmungen,

Bewegungen und Handlungen beschäftigt sind. Vielmehr geht es um den Übergang von individuellen Akten der Beteiligten zu genuin interaktiven Hervorbringungen, mit denen ein zwischen den Anwesenden gemeinsames Hier etabliert wird. Das »Ko-« impliziert deshalb, dass sich die Beteiligten aufeinander beziehen müssen. Das geht nur mit geteilter Aufmerksamkeit, untereinander und aufeinander abgestimmten Bewegungen und praktizierter Zusammenarbeit. Nur auf diese Weise kommen Wahrnehmungen, Bewegungen und Handlungen als Erscheinungsformen kopräsenzbasierter Interaktion zustande. In diesem Sinne geht es um ein soziales Hier als Teil von »Ko-Präsenz«. Als solches bezieht es sich nicht auf ein kognitives Konstrukt, das einzelnen Individuen zugeordnet werden kann, sondern auf ein kommunikatives Konstrukt, das sich an der Oberfläche der Interaktion manifestiert – und dort (und nur dort) auch nachweisbar ist.⁵⁵

Ko-Orientierung, Ko-Ordination und Ko-Operation stehen in einem Verhältnis ansteigender Abhängigkeit voneinander: Ko-Ordination setzt in der Regel Ko-Orientierung voraus, und Ko-Operation setzt ihrerseits in der Regel Ko-Ordination und Ko-Orientierung voraus. Typischerweise fallen der Wahrnehmungs-, Bewegungs- und Handlungsraum empirisch also zusammen (eben als Interaktionsraum), so dass die Unterscheidung vor allem analytischen Wert hat. Das muss aber nicht so sein. Es gibt aufschlussreiche Fälle, in denen der Interaktionsraum nicht zugleich Wahrnehmungs-, Bewegungs- und Handlungsraum ist. Zu diesen Fällen gehört womöglich das, was Goffman als unfokussierte Interaktion bezeichnet hat und worunter er z.B. flüchtige Begegnungen unter Fußgängerinnen und Fußgängern in einer Einkaufszone oder auch das gemeinsame Verweilen auf einem Bahnsteig, in einem Wartezimmer oder einem Fahrstuhl verstanden hat.⁵⁶ Es sind dies zweifellos Erscheinungsformen von Kopräsenz, weil die Beteiligten wahrnehmen können, dass sie wahrgenommen werden. Allerdings verfolgen sie damit nicht automatisch weitere Ziele und kein über das Umgehen mit und Sicheinstellen auf Wahrnehmungswahrnehmung hinausgehendes Engagement. Das ist typischerweise in der Öffentlichkeit der Fall, die gerade dadurch definiert werden könnte, dass Wahrnehmungswahrnehmung unter den öffentlich Anwesenden so gut wie nicht verhindert werden kann, so dass Interaktion mehr oder weniger zwangsläufig anläuft – wie es für das öffentliche Leben in modernen urbanen Kontexten typisch ist.⁵⁷ Natürlich kann sich auf diesem Nährboden jederzeit Weiteres ergeben, aber die Interaktion kann genauso gut auch im reduzierten Stand-by-Modus verbleiben, ohne dass ihr etwas fehlen würde. Es ist von daher nicht ganz befriedigend, wenn man diesen Fall

von Interaktion nur defizitär als *unfokussiert* beschreibt.⁵⁸ Wir schlagen vor, darin Erscheinungsformen von Kopräsenz zu sehen, in denen sich die Situierung tatsächlich dominant auf Ko-Orientierung beschränkt, wobei es immer wieder auch zu Ko-Ordination kommen kann, wenn man z. B. daran denkt, wie Fußgänger:innen einander auf dem Gehweg ausweichen.⁵⁹ Aber es muss eben nicht automatisch auch zu weitergehenden Formen der Ko-Operation kommen, die über das *Vermeiden* von Ko-Operation hinausgehen würden.

Fälle, in denen sich Kopräsenz auf Ko-Orientierung und Ko-Ordination beschränkt, sind nicht nur für den Bereich des öffentlichen Lebens in modernen urbanen Kontexten relevant. Manches spricht dafür, dass Ko-Orientierung und Ko-Ordination auch in solchen Settings besonders relevant sind, in denen sich Formen von Ko-Operation erst noch entwickeln, sei es in phylo- oder in ontogenetischer Hinsicht. Z. B. gibt es Evidenz für Formen von Ko-Orientierung zwischen Erwachsenen und Kleinkindern, die als Schrittmacher für weitergehende Formen von Ko-Ordination gelten können und in und mit denen es überhaupt erst zu rudimentären Formen von Kopräsenz im hier eingeführten Sinn kommt. Ko-Orientierung (und Ko-Ordination) könnten daher Vehikel für Kopräsenz sein, indem sie in der Ontogenese die Voraussetzung dafür schaffen, sozial »anwesend« sein zu können, um auf diese Weise so etwas wie Interaktionsfähigkeit zu erwerben. Aus dieser Blickrichtung, die mit neueren Untersuchungen aus der evolutionären Anthropologie sehr gut kompatibel ist, erscheint die Situierung nicht nur als eine Aufgabe unter anderen Interaktionsaufgaben, sondern als diejenige Aufgabe, die onto- und phylogenetisch am Anfang der Entwicklung von Kopräsenz steht. Ko-Orientierung und Ko-Ordination sind nicht nur nicht auf Sprache angewiesen, sondern auch in weiterer Hinsicht vergleichsweise voraussetzungsarm. So gibt es Evidenz für reduzierte Formen von Ko-Orientierung und Ko-Ordination (z. B. im Sinne des Folgens der Blickrichtung Anderer) ohne reflexive Aufmerksamkeit im Vollsinn sozialer Kopräsenz.⁶⁰ Ko-Orientierung und Ko-Ordination wären dann nicht nur Vehikel für die Etablierung von Kopräsenz, sondern auch mögliche Vorläufer und Vorstufen von Kopräsenz. Es wäre theoretisch gut denkbar, dass Kopräsenz daraus erwächst, dass Ko-Orientierung und Ko-Ordination allmählich als Ergebnis von Ko-Operation mehr und mehr eingespielt, erlebt und behandelt werden. Was wir oben mit Goffman als unfokussierte Interaktion behandelt haben, wäre dann schon ein evolutionär höchst voraussetzungsreicher Fall, weil das Vermeiden von Ko-Operation natürlich ein Bewusstsein und eine Erwartung von Ko-Operation impliziert, wie sie onto- und phylogenetisch gerade noch nicht vorausgesetzt werden

kann und mit Ko-Orientierung und Ko-Ordination auch noch nicht zwingend verbunden war. Im Vergleich zwischen menschlichen und nicht menschlichen Primaten (wie Schimpansen) hat Michael Tomasello diesbezüglich auf die Annahme geteilter Intentionalität als Kriterium für wechselseitig geteilte, reflexive Aufmerksamkeit (im Sinn entwickelter Kopräsenz) hingewiesen, die mit Ko-Orientierung und Ko-Ordination (noch) nicht automatisch mit vorausgesetzt werden kann.⁶¹ Diese Annahme verträgt sich gut mit dem Kommunikationsbegriff der neueren Systemtheorie, für den die Unterstellung der Absicht einer Mitteilung konstitutiv ist. Interaktion und Kopräsenz kommen demnach dadurch zustande, dass die Absicht der Mitteilung des eigenen »Anwesend«-seins wechselseitig unterstellt werden kann.⁶² Kopräsenz ist also nicht voraussetzungslos, was die Unterstellungen und Erwartungen betrifft, die bezüglich der Interaktionsfähigkeit der Beteiligten im Spiel sind. Indem die Interaktionsforschung an dieser Stelle lange Zeit wie selbstverständlich von erwachsenen Menschen als Interaktionsbeteiligten ausgegangen ist, hat sie diese Frage weitgehend außer Acht gelassen. Diese »Agency«-Idealisierung gehört zum Mythos Kopräsenz. Sie hat lange den Weg verstellt, sich darüber Gedanken zu machen, worin die für Kopräsenz grundlegenden Teilnahmebedingungen bestehen und wie sie erworben worden sind (phylogenetisch) und nach wie vor erworben werden (ontogenetisch). Hier liegt zugleich ein Schlüssel zum Einblick in Phänomene des Wandels von Kopräsenz, wenn man daran denkt, wie Interaktionsfähigkeit (medien-)technisch suggeriert werden kann.⁶³

Wie immer man die entwicklungs- und evolutionstheoretische Fruchtbarkeit der Unterscheidung von Ko-Orientierung, Ko-Ordination und Ko-Operation auch einschätzen mag, zeigt sich darin doch recht anschaulich, wie ausgehend von grundlegenden Problemen der Situierung der Interaktionsraum vom Wahrnehmungs- und Bewegungsraum zum Handlungsraum an Profil gewonnen hat. Ko-Operation zeichnet sich im Vergleich mit Ko-Orientierung und Ko-Ordination vor allem dadurch aus, dass Ko-Orientierung und Ko-Ordination von vornherein auf einen bestimmten Typ von Miteinander (»withness«), auf ein gemeinsames interaktiv zu bearbeitendes Projekt (»joint project«) und einen darauf bezogenen Aktivitätstyp bezogen sind. Das kommt bereits in den typischen Kopräsenzkonfigurationen zum Ausdruck, die wir bereits besprochen haben (»face-to-face«, »side-by-side«, »face-to-back«) und die bis in die feinen Verästelungen der ausdifferenzierten sozialen Praktiken des Ko-Operierens in unserer modernen Gesellschaft reichen. Das Sehen einer Konfiguration von Anwesenden ermöglicht es uns deshalb

in vielen Fällen, auf das fragliche interaktive Projekt zurückzuschließen. Es ist schließlich im Übergang von Ko-Orientierung und Ko-Ordination zu Ko-Operation zu lokalisieren, dass die Sequentialität als Ordnungsprinzip der Interaktion an Relevanz gewinnt. Die Gleichzeitigkeit von Wahrnehmungen und Bewegungen, wie sie in Prozessen der Synchronisation zum Ausdruck kommt, wird überführt in das geordnete Nacheinander und die Reihenfolge interaktiver Züge – mit den bekannten Implikationen für das Sichabwechseln als Folgeproblem. Nicht zufällig stoßen wir an dieser Stelle auf das Sprechen-und-Zuhören. Es ist vielleicht der spektakulärste Ausdruck des Übergangs zur Ko-Operation: Die Sprachspiele, in und mit denen wir Dinge mit Worten tun und »sprachlich handeln«, sind als »Gespräch« Ausdruck und Faktor einer weitgehenden Verselbständigung von Ko-Operation gegenüber Ko-Orientierung und Ko-Ordination. Das bedeutet nicht, dass Ko-Operation nicht auch weiterhin auf Ko-Orientierung und Ko-Ordination angewiesen wäre. Aber die Ko-Operation verlagert sich mit dem Sprechen-und-Zuhören auf den Austausch von Worten, die sich mehr und mehr von ihrer Anbindung an eine konkrete »Sprechsituation« lösen. Entsprechend treten die Ko-Orientierung und die Ko-Ordination in den Hintergrund – und in der Sprache wird die Situationsbindung des Sprechens-und-Zuhörens zu einem Spezialphänomen (Zeigfeld, Deixis). Interaktion wird mehr und mehr zur *verbalen* Interaktion. Unsere These ist, dass damit Folgeprobleme der (verbalen) Interaktion entstehen, die sich evolutionär aus der erfolgreichen Bearbeitung des Problems der Situierung ergeben (haben). Es sind dies Folgeprobleme wie das schon angesprochene Sichabwechseln, der Bezug auf Welt und das Beitragen zu Themen, die Selbst- und Fremddarstellung, die Rahmung und schließlich auch die Aufwertung von Kopräsenz in Form von markierten Episoden mit Anfang und Ende. So gesehen wäre die Situierung und hier speziell der Übergang zur Ko-Operation die Keimzelle für verbale Interaktion, wie wir sie heute kennen und beschreiben (s. Abb. 14).

Diese Darstellung gibt zunächst der Situierung, d.h. der Schaffung eines Interaktionsraums, ein besonderes Gewicht. Das hat damit zu tun, dass die Situierung mit der Ko-Orientierung unmittelbar an der für Kopräsenz zentralen Konstitution von reflexiver Wahrnehmung und geteilter Aufmerksamkeit beteiligt ist. Weiter gibt sie speziell der Ko-Operation, d.h. der Schaffung eines Handlungsraums, besonderes Gewicht für die Onto- und Phylogenese von Kopräsenz. Wie aus einer Keimzelle haben sich, so die These, aus den Anfängen der mit Ko-Operation verbundenen gemeinsamen Projekte Folgeprobleme entwickelt und profiliert, wie wir sie heute mit Kopräsenz verbinden.

Die problemorientierte Profilierung von Kopräsens hat mit dem Einsatz des Sprechens-und-Zuhörens die Evolution von Sprache, in der Form, wie sie uns heute geläufig ist, maßgeblich vorangetrieben. Diese Sichtweise trifft sich vielleicht nicht zufällig mit der herausragenden Rolle, die der Ko-Operation nicht nur für die Evolution von Sprache, sondern für die Menschwerdung generell in neueren Arbeiten aus der evolutionären Anthropologie zugesprochen wird.⁶⁴ Wir wollen auf diese Dimension der Evolution von Kopräsens hier nicht weiter eingehen und stattdessen auf die Unwahrscheinlichkeit der Interaktion zurückkommen und einen abschließenden Blick auf die Ressourcen werfen, mithilfe derer diese Unwahrscheinlichkeit der Situierung tagtäglich überwunden werden kann.

Abb. 14: Folgeprobleme von Ko-Operation (eigene Darstellung)



III.

Das Problem der Situierung begründet eine grundlegende Unwahrscheinlichkeit der Interaktion, die noch vor den Unwahrscheinlichkeitsschwellen der Kommunikation liegt, die Niklas Luhmann in seinem Konzept zur »Unwahrscheinlichkeit der Kommunikation« ansetzt.⁶⁵ Dort beginnt die Unwahrscheinlichkeit der Kommunikation mit dem »Verstehen«, und sie wird in dieser Ausprägung nach Luhmann maßgeblich durch die evolutionäre Errungenschaft des Mediums Sprache überwunden. Wir haben es hier also bereits mit der Unwahrscheinlichkeit der *verbalen* Interaktion zu tun. Die Unwahrscheinlichkeit der Situierung ergibt sich aber noch vor dem Sprechen- und -Zuhören schon daraus, dass ein gemeinsam geteilter Interaktionsraum etabliert werden muss. Das fängt mit dem Wahrnehmungs- und Bewegungsraum an und endet beim Handlungsraum. Jeder dieser Räume begründet eine eigene Problematik (Unwahrscheinlichkeit). Wo etwa fängt der Wahrnehmungsraum an und wo hört er auf? Wie kann ich sicherstellen, dass mein Gegenüber das Gleiche wahrnimmt und als relevant erachtet, was ich gerade wahrnehme und als relevant erachte? Gegeben die Perspektivität der jeweils individuellen Wahrnehmung ist es unwahrscheinlich, dass die Beteiligten wie selbstverständlich von einer gemeinsam geteilten Situation ausgehen. Mein Gegenüber sieht z.B. aus prinzipiellen Gründen und unaufhebbar etwas anderes, als ich gerade sehe – und er oder sie kann auch nicht sicher sein, was genau ich gerade wahrnehme, weil es vielleicht nur für mich gerade überhaupt wahrnehmbar und/oder relevant ist. Gut kann man sich vorstellen, wie die Beteiligten daran scheitern würden, wenn sie sich vor jeder Interaktionsepisode untereinander zunächst ausdrücklich darüber verständigen (und »aushandeln«) wollten, was an Wahrnehmungen bis auf weiteres als geteilt anzusehen ist. Hinzu kommt, dass unsere Wahrnehmungen nicht ein für alle Male fixiert, sondern höchst dynamisch sind. Menschen sind mobile Sensoren, sie bewegen sich und mit ihren Bewegungen verändern sich die Wahrnehmungen. Es stellt sich deshalb die Frage, wie der Wahrnehmungs- in den Bewegungsraum überführt werden kann und wie bei mobilen Beteiligten überhaupt sichergestellt werden kann, dass sie »zusammen« bleiben, und was genau das heißt mit Bezug auf Nähe und Distanz, Neben-, Hinter- und Miteinander. Es ist leicht zu sehen, dass Bewegung geteilte Wahrnehmung gefährdet; für die, die um einen Tisch herumsitzen (bleiben), ist es leichter, »zusammen« zu bleiben, als für die, die zusammen einen Spaziergang machen oder ein Museum besuchen. Müssen sich die Beteiligten also fortwährend verständigen und sich

im Grunde jederzeit aufs Neue explizit vergewissern, ob sie noch über einen geteilten Wahrnehmungs- und Bewegungsraum verfügen und wie ihre Wahrnehmungen und Bewegungen auf das jeweilige Handlungsziel zugeschnitten und angepasst werden können? Offenkundig kann es so nicht gehen, aber wie geht es dann? Diese Frage führt uns zu den Ressourcen der Interaktion, die sich bei den Anforderungen der Ko-Orientierung, Ko-Ordination und Ko-Operation besonders anschaulich zeigen.

Wir haben schon betont, dass Ko-Orientierung typischerweise im Medium reflexiver Wahrnehmung und entsprechend wechselseitig geteilter Aufmerksamkeit erfolgt. Deixis, nicht nur sprachlicher, sondern grundsätzlich multimodaler Natur, ist ein prototypisches Mittel zur Bearbeitung der Aufgabe der Ko-Orientierung. Mit ihr wird eine konkrete Wahrnehmung selbst hör- und sichtbar, d.h. wahrnehmbar als Teil des gerade relevanten Wahrnehmungsraums. Entgegen dem Sprachgebrauch muss man betonen, dass deiktische Ausdrücke und Gesten nicht »zeigen«, sondern die Aufmerksamkeit auf einen Wahrnehmungsakt, typischerweise einen Blick auf etwas in der Umgebung, lenken. Gesten sind, vereinfacht gesagt, visualisierte Blicke. Sie schalten um auf einen anderen »Kanal« und wollen die Partnerin oder den Partner wissen lassen, dass es etwas zu sehen gibt, was sich zu sehen lohnt. Sie kommunizieren also die Absicht einer Mitteilung. In dieser Etablierung einer Unterscheidung von Information (»da« ist etwas) und Mitteilung (von dem ich will, dass Du es auch siehst) liegt ihr evolutionäres Potential für Kommunikation (»joint attention«) und ein Schlüssel für die weitere Entwicklung in Richtung Ko-Operation.⁶⁶

In dem Maße, in dem der Interaktionsraum als Bewegungsraum durch Dynamik, Flexibilität und Fluidität ausgezeichnet ist, ist seine Etablierung über die Deixis hinaus auf die weitere Ausstattung beweglicher Sensoren und in Bewegung begriffener Sensoren angewiesen. Kopräsens kann dazu in vielen Fällen auf die Sensorik und Lokomotorik des Menschen zurückgreifen. Dazu gehört ein komplexes Zusammenspiel von Kinesik⁶⁷ und Proxemik,⁶⁸ das sich z.B. in so etwas wie Rudel- und Herdenverhalten (»pack behavior«) zeigt,⁶⁹ aber grundsätzlich das System wechselseitig abgestimmter Bewegungen im Raum meint, mit dem Bewegungen nicht nur reguliert, sondern auch in ihrer für die Andere oder den Anderen relevanten Dimension verdeutlicht werden. Während der Bewegung auf verlässliche und nicht zufällige Weise »synchron« in Reich- und Hörweite zu verbleiben, ist insofern ein Mittel, das Problem des Zusammenbleibens zu lösen, ohne dass darüber gesprochen werden muss. Synchronisation ist nicht nur ein effektives, sondern auch anspruchs-

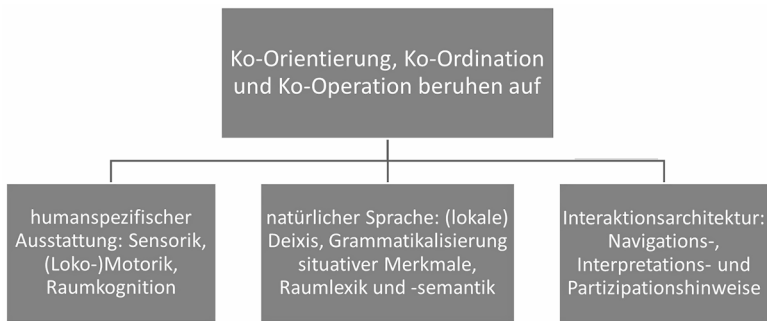
volles Mittel, das in Spezialfällen der Etablierung von Bewegungsräumen (= Ko-Ordination) spezielle Heraus- und Anforderungen an die Beteiligten stellt. Der Tanz ist dafür ein gerne zitiertes Beispiel.⁷⁰ Er leitet zugleich über zum Problem der Ko-Operation, weil er zeigt, dass und wie Wahrnehmung und Bewegung Teil einer bestimmten sozialen Praxis sind. Typischerweise hinterlässt diese Praxis auch im Raum selbst ihre Spuren in Form von Interaktionsarchitektur.⁷¹ Ein sehr illustratives Beispiel dafür ist der Sport, hier speziell der Bereich der Mannschaftssportarten, die von der Bewegung im Raum leben und dafür auf besonderen »Plätzen« ausgetragen werden. Man denke an ein Fußball- oder Handballfeld mit seinen klaren Begrenzungen und Markierungen in Form von Linien, Feldern und Behältnissen, um einen Eindruck davon zu bekommen, wie ein konkreter Handlungsraum zum Spielraum wird und wie dieses hoch unwahrscheinliche Handlungsspiel durch die Interaktionsarchitekturen der Spielfelder möglich und wahrscheinlich gemacht werden kann. Ein Blick in eine Mehrzweckhalle erhellt, wie komplex die Markierungen sein können, mit denen der Handlungsraum zum Spielraum wird (s. Abb. 15).

Abb. 15: Handlungsraum als Spielraum (Quelle: ASVZ-Archiv der ETH Zürich)



Zugleich zeigt die Abbildung, dass solche Markierungen auf die Spielregeln für die fragliche soziale Praxis verweisen, wobei sich die Markierungen in diesem Fall überlagern und auf eine für die Laien oder den Laien verwirrende Weise auf unterschiedliche Regeln und unterschiedliche Spiele mit unterschiedlichen Feldern hinweisen. Solche »Benutzbarkeitshinweise«⁷² sind offenkundig stark wissens- und vertrautheitsabhängig. So künstlich dieses Beispiel wirkt, so anschaulich demonstriert es, wie durch die Gestaltung des Raumes (hier u.a.: durch die Markierung des Fußbodens) Räume für Wahrnehmung, Bewegung und Handlung geschaffen werden. Die, die in dieser Halle spielen wollen, können sich darauf wie selbstverständlich beziehen, indem sie bestimmte Positionen einnehmen, bestimmte Felder tunlichst nicht betreten und/oder bestimmten Zonen besondere Beachtung schenken.⁷³ Weiter zeigt das Beispiel, wie sehr Kopräsens durch die Interaktionsarchitektur konfiguriert werden kann als ein Beisammensein unter den Vorgaben des Raumes. Natürlich kann keine Interaktionsarchitektur verhindern, dass die Beteiligten dort etwas machen, was vom Raum nicht vorgesehen ist (z.B. auf dem Hallenboden Kaffee trinken). Aber es ist auch klar, dass die Beteiligten in diesem Fall nicht die Ressourcen nutzen, die der Raum für eine bestimmte soziale Praxis bereitstellt und die er damit möglich und wahrscheinlich macht. Speziell die Welt der institutionalisierten Interaktion ist voll von solchen architektonischen Ressourcen, die prinzipiell nicht anders als in der Sporthalle auf die »Spielregeln« hinweisen, die für die fragliche soziale Praxis Geltung beanspruchen. Man denke nur an das Beispiel des Gerichtssaals mit seiner differenzierten Aufteilung von Plätzen, die als Positionen im Raum zugleich soziale Positionen sind, in denen Zuständigkeiten, Befugnisse und Verpflichtungen vertrautheitsabhängig zum Ausdruck kommen.⁷⁴

Die Beispiele für Ko-Orientierung, Ko-Ordination und Ko-Operation zeigen, nicht überraschend, die schon eingeführten rezenten Ressourcen der Interaktion: Körperlichkeit, Sprache und Architektur, wie sie sich mit Bezug auf die Gegenwart entwickelt haben.⁷⁵ In schematischer Vereinfachung lassen sie sich mit Bezug auf das übergreifende Problem der Situierung wie in Abb. 16 dargelegt zusammenfassen.

Abb. 16: *Rezente Ressourcen der Situierung (eigene Darstellung)*

Die Abbildung spezifiziert die Ressourcenqualität von Körper, Sprache und Architektur für das Problem der Situierung. Wie ausgeführt kann die Ko-Orientierung, Ko-Ordination und Ko-Operation auf den menschlichen Körper als einen hoch entwickelten, mobilen und intelligenten Sensor für die Navigation im Wahrnehmungs- und Bewegungsraum zurückgreifen. Die Lösung dieser Interaktionsprobleme manifestiert sich entsprechend im Blicken, Zeigen, Bewegen, Annähern und Abstandhalten. Die humanspezifische Ausstattung des Menschen ist auch die Antwort auf diese Probleme der Interaktion. Das alles geht auch ohne Sprache – aber es geht sehr effektiv *mit* Sprache, wenn sie evolutionär zur Verfügung steht. Wenn sich die Ko-Operation zwischen den Beteiligten in Sprachspielen manifestiert, also die Form des Sprechens-und-Zuhörens annimmt, ist auch die Sprache sehr effektiv an der Herstellung des Interaktionsraums beteiligt, indem sie spezifische Leistungen erbringt. Charakteristisch ist deshalb die Ausnutzung körperlicher *und* sprachlicher Ressourcen. Das zeigt sich z.B. in der Multimodalität lokaler Deixis, die auf Gesten, Blicken, Körperzuwendungen und deiktischen Ausdrücken beruht und so sprachliche Raumreferenz mit der Konfiguration der Körper im Raum verschränkt, so dass man ohne weiteres von verkörperter Raumreferenz sprechen kann (»embodiment«, »intercorporeality«).⁷⁶ Das gilt auch und gerade für die Grammatikalisierung situationsbezogener Merkmale, wie sie in Ausdrücken wie »hier« und »da«, »oben« und »unten«, »links« und »rechts«, »vorne« und »hinten«, »hin« und »weg« zum Ausdruck kommt. Auch ohne dass solche Ausdrücke Wahrnehmung unmittelbar adressieren müssen (wie das bei »starker« bzw. »echter« Deixis der Fall ist), beziehen sie sich auf den Interaktionsraum, wie er als »konkret-anschaulicher (<deiktischer>) Kommunikationsraum [...] durch die leibliche Konfiguration der miteinander

sprechenden Personen gebildet ist«. ⁷⁷ Mit ihnen haben wiederkehrende Situierungsanforderungen eine sprachliche Sedimentierung gefunden, so dass man von einer Grammatikalisierung vielfach bewährter Navigationsparameter sprechen könnte. Das gilt z.B. für die »Dimensionen der Leiblichkeit«, die im Deutschen in den Lokaladverbien als »Vertikalität« (oben/unten), »Frontalität« (vorne/hinten), »Lateralität« (rechts/links) und »Interiorität« (innen/außen) zum Ausdruck kommen. ⁷⁸ Grammatikalisierungen wie diese setzen deshalb (nur noch) ein Minimum geteilter Aufmerksamkeit voraus, auf dessen Grundlage beim Sprechen-und-Zuhören der Interaktionsraum im Hintergrund der Interaktion sprachlich aufscheinen kann.

Oftmals wird die Situierung im Hier und Jetzt durch den menschlichen Körper im Vollsinne seiner sensomotorischen Qualitäten »geleistet«, typischerweise auch ohne dass – oder bevor – gesprochen wird. Wenn gesprochen und auf Raum in irgendeiner Weise Bezug genommen wird, lässt sich zeigen, wie dieses Sprechen-und-Zuhören mit der körperlichen Konfiguration der Beteiligten unmittelbar verknüpft ist. Dieses komplexe Zusammenspiel von körperlichen und sprachlichen Ressourcen ist vielfältig beschrieben worden. Was dagegen in diesem Zusammenhang bis heute noch weitgehend außer Acht gelassen wurde, ist die Ressourcenqualität des Raumes selbst, wie sie beispielsweise in den o. illustrierten Markierungen von Spielfeldern in der Sporthalle zum Ausdruck kommt (s.o. Abb. 15). Wir schlagen vor, diese Ressourcenqualität des Raumes als Interaktionsarchitektur zu verstehen, die sich in architektonischen Benutzbarkeitshinweisen manifestiert. Im Gegensatz zum Interaktionsraum, der auf Anwesenheit beruht und ein Teil von Kopräsenz und deshalb emergent, dynamisch, flexibel und fluide ist, beruht die Interaktionsarchitektur als Ausdruck von Kommunikation auf Benutzbarkeit. Sie liefert Anknüpfungspunkte für Kopräsenz und macht insbesondere Prozesse der Ko-Orientierung, Ko-Ordination und Ko-Operation hochgradig möglich und wahrscheinlich – ohne dass darüber gesprochen werden muss, »einfach« dadurch, dass die Beteiligten Positionen im Raum einnehmen und damit Navigations-, Interpretations- und Partizipationshinweise relevant setzen. ⁷⁹ Mit der Konfiguration von Kopräsenz durch Interaktionsarchitektur werden entsprechend höchst anspruchsvolle Formen der Situierung zu einer alltäglich-routinisiert vollzogenen Selbstverständlichkeit, an die u.a. das Sprechen-und-Zuhören unmittelbar anknüpfen kann. Welche Folgeprobleme verbaler Interaktion über die Situierung hinaus damit ins Spiel kommen, wollen wir im Folgenden erläutern.

Mehr als nur Sprecherwechsel: Aufgaben in der verbalen Interaktion

So abstrakt das theoretische Axiom von der Unwahrscheinlichkeit der Interaktion ist, so konkret und unmittelbar empirisch fruchtbar sind die Beobachtungen, die man auf der Grundlage dieses Axioms machen kann. Entsprechend findet sich der Grundgedanke von Interaktion als einem aufgabenlösenden Problemzusammenhang in vielen interaktionsanalytischen Ansätzen aus den 1970er Jahren, besonders prominent z.B. in der »Rahmenanalyse« von Erving Goffman.⁸⁰ Markant hervor tritt die Idee der interaktiv zu bearbeitenden Aufgaben (»jobs«) in den frühen Aufsätzen der Konversationsanalyse, wenn es darum geht, »Probleme« und Fragen zu explizieren, als deren »Lösungen« und Antworten die dokumentierbaren sprachlichen Erscheinungsformen der Interaktion entdeckend rekonstruiert werden können. In der Frühphase der deutschsprachigen Rezeption der Konversationsanalyse hat dieser Gedanke in einem viel zitierten Aufsatz von Werner Kallmeyer und Fritz Schütze eine gewisse Prominenz erlangt.⁸¹ Kallmeyer und Schütze hatten in diesem Aufsatz, der in den 1970er Jahren viel dazu beigetragen hat, die Konversationsanalyse in der deutschsprachigen Linguistik bekannt zu machen (und der zugleich ein eindrucksvolles Dokument für ein kongeniales Miteinander von Linguistik und Soziologie ist), die Idee entwickelt, dass sich Interaktion auf eine theoretische Systematik von konstitutionslogischen Anforderungen zurückführen lassen müsste. Mit diesem Konzept, das über die Rezeption der Konversationsanalyse bereits hinausging, hatten die Autoren allerdings dem empirischen und methodologischen Stand der Forschung sehr weitgehend vorgegriffen.⁸² Knapp 50 Jahre später, vom Erscheinen des genannten Aufsatzes an gerechnet, lässt sich der seinerzeit erhobene Anspruch, interaktionskonstitutive Probleme auf eine systematische Weise zu explizieren, zumindest in Ansätzen und empirisch gut fundiert einlösen. Wir schlagen dazu eine Systematik von sechs konstitutiven Aufgaben vor (Interaktionseröffnung und -beendigung, Sprecherwechsel, Themenorganisation, Selbst- und Fremddarstellung, Rahmung und Situierung), die sich auf grundlegende Fragen und Probleme der Interaktion beziehen lassen, und skizzieren zunächst allgemeine Merkmale dieser Aufgaben. Im Anschluss gehen wir dann auf jede Aufgabe näher ein, wobei wir uns auf eine allgemeine Darstellung beschränken, ohne Aufgabe für Aufgabe in die Details zu gehen. Wir wollen an dieser Stelle keine weitere Einführung in die Konversationsanalyse schreiben, sehen es aber als eine Art von Bestätigung an, dass man entlang dieser Aufgaben und ihrer in der Sekundärliteratur inzwischen zahlreich beschriebenen Lösungsverfahren

ohne weiteres eine solche Einführung schreiben könnte.⁸³ Stattdessen greifen wir neben der allgemeinen Charakteristik ein paar Punkte heraus, die uns vor dem Hintergrund unserer Fokussierung auf Kopräsenz und angesichts des aktuellen Forschungsstandes besonders relevant erscheinen.

I.

Wie einleitend ausgeführt, stellen wir uns Interaktion als Bearbeitung der mit Kopräsenz emergierenden Problemstellungen vor. Wenn man Anwesenheit nicht als äußerliche Bedingung von Kopräsenz auffasst, sondern als eine Errungenschaft, ist die *Herstellung von Anwesenheit* die erste und grundlegende interaktive Leistung, als deren Entfaltung sich alle weiteren Aufgaben bis in die feinsten Verästelungen moderner Interaktionssysteme verstehen lassen. Die Differenzierung interaktiver Aufgaben lässt sich also als schrittweise Entfaltung der in Kopräsenz angelegten Problemstellungen begreifen, wobei wir nicht vergessen dürfen, dass wir in unserer Darstellung zumeist von gegenwärtigen Erscheinungsformen von Kopräsenz ausgehen, die z.B. die Konfiguration durch Sprache bereits voraussetzen. Von diesem Bias können wir unsere Darstellung nicht (ganz) befreien. Aber immerhin sind wir in unserem Modell auf eine Aufgabe gestoßen, mit der ein Fenster aufgeht und den Blick auf frühere, weniger anspruchsvolle Entfaltungen von Kopräsenz freigibt. Das ist die Aufgabe der Situierung, die wir deshalb vorausgehend in einem eigenen Unterabschnitt behandelt haben.

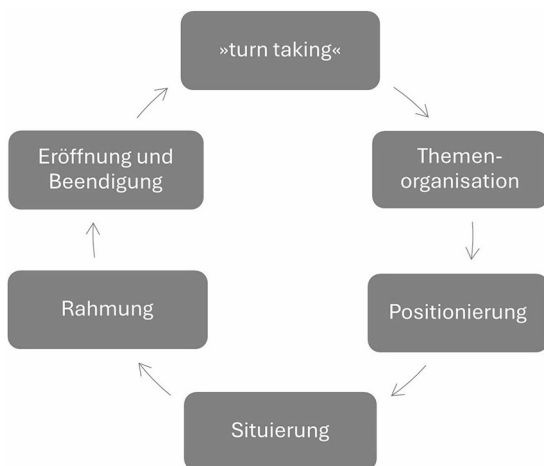
Wenn man keine Angst vor Vereinfachungen hat, lassen sich die grundlegenden Probleme der Interaktion auf einen Zusammenhang von sechs Fragebereichen bringen. Die folgende Übersicht (s. Abb. 17) veranschaulicht diesen Zusammenhang, aus dem heraus die einzelnen Fragen jeweils auf ein in und mit Interaktion zu lösendes Problem hinweisen. Darin spiegelt sich nicht zufällig die neuere Forschungsgeschichte der Konversationsanalyse und Interaktionslinguistik.⁸⁴

Abb. 17: Die Unwahrscheinlichkeit der Interaktion (eigene Darstellung)



Die hier aufgeführten Fragen sind so gemeint, dass sie auf Probleme im Gegenstandsbereich selbst verweisen – also auf Probleme, die in und mit Interaktion (vereinfacht gesagt: durch die, die durch Teilnahme an Kopräsenz zu Interaktionsbeteiligten werden) irgendwie gelöst werden müssen, damit Interaktion überhaupt in Gang kommen kann. Vor dem Hintergrund dieser Fragen zeigt sich, worin die Schwellen der Entmutigung im Fall von Interaktion bestehen und dass Interaktion tatsächlich ein unwahrscheinliches Unterfangen ist, weil eine Reihe von Schwellen überwunden werden muss, die, vereinfacht gesagt, mit dem Anfangen und Aufhören, dem Drankommen und Beitragen, dem Sichpositionieren mit und gegenüber Anderen, dem Sichzurechtfinden in einer Situation und dem Sicheinordnen in einen Interaktionskontext zu tun haben. Die analytische Aufgabe besteht dann darin zu rekonstruieren, wie diese Fragen interaktiv beantwortet werden. Es geht dabei um Probleme, so die These, die so grundlegend sind, dass sie in jeder x-beliebigen Interaktion auf irgendeine Weise gelöst werden müssen. Von daher handelt es sich um Anforderungen an die Interaktion, die in der einschlägigen Forschung (meistenteils) gut bekannt sind. Die Übersicht in Abb. 18 weist die Aufgaben der Interaktion entsprechend durch (meistenteils) gut bekannte Stichworte aus.

Abb. 18: Interaktion als Problemlösung (eigene Darstellung)



Eröffnung und Beendigung stehen dafür, dass sowohl der Beginn als auch der Abschluss der Interaktion nicht extern vorausgesetzt, sondern mit Bordmitteln der Interaktion geleistet werden müssen. Anfang und Ende gehören zur Interaktion dazu. »Turn taking« verweist darauf, dass in dem Maße, in dem gesprochen-und-zugehört wird, die Beteiligten sich als Sprecher:in und Hörer:in abwechseln müssen. Themenorganisation meint, dass geklärt werden muss, worüber zu sprechen ist und wie einzelne Beiträge sich zu einem gemeinsamen Thema verbinden (lassen). Positionierung hat damit zu tun, dass es, etwas altmodisch gesagt, neben dem Inhalts- auch einen Beziehungsaspekt der Kommunikation gibt, unter dem Selbst- und Fremddarstellungen kommuniziert werden müssen. Situierung bezieht sich darauf, dass die laufende Interaktion in der sinnlich wahrnehmbaren Umgebung irgendwie verankert werden muss. Schließlich deutet Rahmung darauf, dass der gesellschaftliche Kontext in den für die laufende Interaktion relevanten Aspekten bestimmt werden muss. Die hier benutzten Stichworte sind in der einschlägigen Forschung unterschiedlich gut etabliert, und auch die dahinter stehenden Aufgaben können als unterschiedlich gut erforscht gelten. Sehr prominent ist die Aufgabe des *Sichabwechselns* mit ihrer »Who comes next«-Orientierung (»turn taking«, Sprecherwechsel), die bis heute als *die* Entdeckung der Konversationsanalyse gelten kann. Die Konfiguration von Kopräsenz durch Sprache macht sich bei diesem Problem besonders deutlich bemerkbar: Es steht stellvertre-

tend für die Art von Reihenfolge- und Nacheinanderproblemen, die ko-evolutionär mit Sprache (und den Vorläufern von Sprache) in die Kopräsenzsituation gekommen sind und ihrerseits die Evolution von Sprache massiv herausgefordert haben dürften. Auch die *Eröffnung und Beendigung* können zu den klassischen Aufgabenbereichen zählen, die von der Konversationsanalyse früh untersucht worden sind (mit Einschränkungen auch die *Themenorganisation*). Dagegen ist die Aufmerksamkeit für das Problem der *Rahmung* (und das der damit eng verwandten »Kontextualisierung«) eher über die Soziolinguistik und die Ethnographie der Kommunikation in die Interaktionslinguistik gekommen.⁸⁵ Erst später, dann aber sehr intensiv ist schließlich auch die *Positionierung* (als Oberbegriff für die Selbst- und Fremddarstellung) wieder in den Fokus gerückt, was speziell für den Bereich der sozialen Kategorisierung⁸⁶ gilt, nachdem diese über lange Jahre im Schatten des Sprecherwechsels stand. Schließlich findet sich in diesem Katalog auch die Aufgabe der *Situierung*, die über lange Jahre vergleichsweise wenig erforscht worden ist und erst in den letzten ca. 15 Jahren mit dem Aufschwung der videobasierten Interaktionsforschung vermehrt untersucht worden ist.⁸⁷ Es ist diejenige Aufgabe, die vielleicht am wenigstens durch Sprache geprägt ist, wie wir bereits betont haben.

Bevor wir auf die Aufgaben im Einzelnen eingehen, seien ein paar allgemeine Charakteristika vorausgeschickt. Wie schon einleitend angemerkt, adressieren diese Aufgaben nicht die Beteiligten und ihre psychischen Systeme, sondern das Interaktionssystem, in einer nicht ganz strengen Terminologie: die Beteiligten als Interaktionsteilnehmer:innen. Die Erledigung der Anforderungen verlangt deshalb weder das Bewusstsein noch die Reflexion der Beteiligten; es gehört zu dem, was sie routinemäßig erledigen, ohne darüber ins Grübeln zu kommen: Man sieht jemanden und grüßt, man stellt dem Gegenüber eine Frage (und bringt ihn damit zum Zug), man spricht ein Thema an, das man schon lange anschneiden wollte, man macht einen Vorwurf und stellt sich selbst damit in ein positives moralisches Licht, man zeigt auf einen Gegenstand im Raum, weil sich dort etwas Komisches bewegt hat, man wählt eine förmliche Ansprache, um dem formellen Charakter der Begegnung gerecht zu werden. Die Aufgaben, die mit solchen Aktivitäten bearbeitet werden, bleiben typischerweise verdeckt, und sie kommen nicht zur Sprache. Schon die Vorstellung eines ständigen Erledigens von Aufgaben ist im Alltag kontraintuitiv. Dessen ungeachtet wollen wir hier davon ausgehen, dass die fraglichen Aufgaben zum einen nicht ein für alle Male erledigt werden, sondern immer wieder von Neuem anfallen, und dass zum anderen die Aufgaben nicht nacheinander, sondern gleichzeitig erledigt werden. In-

teraktion ist so gesehen ein ständig arbeitendes »multitasking system«, in und mit dem fortlaufend die genannten Aufgaben (in ihren unzähligen Verfeinerungen und Verästelungen) bearbeitet werden. Jede Erscheinungsform der Interaktion ist manifester Ausdruck dieser Form von Problemlösung. Wann immer Interaktion anläuft, so die These, sind die genannten Aufgaben ausnahmslos im Spiel. Die Erledigung der Aufgaben kann sich im Falle sehr kurzer, flüchtiger Begegnungen, die vielleicht aus nicht viel mehr als einem etwas längeren Blickkontakt bestehen, auf ein Minimum beschränken, ohne dass die Interaktion in allen ihren Aufgabenbereichen überhaupt angelaufen ist – und gerade dadurch ihre Charakteristik als flüchtige und vielleicht sogar fragwürdige Begegnung erhalten. Sobald die Schwelle zur Verbalität überschritten wird, kann man davon ausgehen, dass der gesamte Apparat in Gang kommt.

Auch wenn alle Aufgaben stets gleichzeitig im Spiel sind, sind nicht alle Aufgaben in jedem Moment gleichermaßen relevant und fokussiert. Ein Effekt von Multitasking besteht vielmehr darin, dass die Komplexität von Kopräsens zunimmt und deshalb ein erhöhter Fokussierungsbedarf entsteht im Hinblick auf das, was interaktiv gerade besonders dringlich zu erledigen ist, d.h. welche Aufgabe gerade interaktiv im Vordergrund steht. In der Interaktion geschieht alles, was überhaupt geschehen kann, unter der Bedingung der Wahrnehmungswahrnehmung, vereinfacht gesagt: der Bedingung wechselseitig geteilter Aufmerksamkeit.⁸⁸ Wenn man dieses für Kopräsens unabdingbare Minimum wechselseitig geteilter Aufmerksamkeit *Interaktionsfokus* nennt, ist alles, was dazu beiträgt, diesen Interaktionsfokus zu etablieren, zu verschieben oder sonst wie zu verändern, Ausdruck von Fokussierung. In diesem weiten Sinn ist auch Fokussierung eine konstitutive Anforderung der Interaktion, die immer anfällt und immer irgendwie gelöst werden muss: Sobald Interaktion über eine flüchtige, gleich wieder aufgelöste Begegnung hinausgeht, entsteht ein Interaktionsprofil, das zwischen einem Minimum und einem Maximum Unterscheidungen von Dringlichkeit ausdrückt. Das, was in der Interaktion am dringlichsten zu erledigen ist, steht im auffälligen Vordergrund, das was am wenigsten dringlich ist, im unauffälligen Hintergrund der Interaktion. Fokussierungen dienen dazu, dieses Profil von Vorder- und Hintergrund zu entwickeln, aufrechtzuerhalten und zu modifizieren. Sie sind notwendig, damit Anschlussbedingungen für nächste interaktive Züge wirksam werden können und das Nacheinander interaktiver Züge, insbesondere von Redebeiträgen, als organisierte Reihenfolge erlebt und behandelt werden kann.

Bindet man Fokussierungen auf diese Weise an Probleme der Interaktionsfortsetzung, rücken unter dem Aspekt der Fokussierung sofort *konditionelle Relevanzen* in den Blickpunkt. Das sind interaktive Zugzwänge, mit denen Erwartungen manifest und damit zu strukturell wirksamen Erwartungserwartungen⁸⁹ werden. Alles, was dazu beiträgt, konditionelle Relevanzen zu etablieren, trägt so gesehen zur Fokussierung bei. Das, was durch ein Maximum an konditionellen Relevanzen als maximal dringlich signalisiert wird, ist zugleich das, was im auffälligen Vordergrund der Interaktion steht; das, was als minimal dringlich signalisiert wird, ist zugleich das, was im unauffälligen Hintergrund der Interaktion steht. Das Paradebeispiel für maximale Dringlichkeit sind lokale Zugzwänge, wie sie innerhalb von Paarsequenzen auftreten. So kann durch einen verbalen Gruß etwa ein vielleicht noch labiler, unsicherer Wahrnehmungsfokus durch die Etablierung der Erwartbarkeit der Erwartung des Zurückgrüßens stabilisiert werden und auf diese Weise die Aufgabe der Interaktionseröffnung als maximal dringlich in den Vordergrund der Interaktion rücken. Das, was als dringlich und vordergründig gelten soll, muss also durch starke konditionelle Relevanzen ausgestattet werden, die eindeutige Reihenfolgeerwartbarkeiten etablieren. Konditionelle Relevanzen sind nach dieser Auffassung der genuin interaktive Ausdruck von Fokussierungen. Nicht immer freilich muss dieser Zusammenhang in der Interaktion so klar und deutlich hervortreten wie im Falle von Paarsequenzen. Die analytische Herausforderung erwächst deshalb daraus zu zeigen, wie alles, was gesprochen und gehört wird, grundsätzlich dazu beiträgt, Erwartungen an das Nächste zu manifestieren und damit als solche erwartbar zu machen.⁹⁰ In diesen Prozess der Fokussierung ist die Erledigung der konstitutiven Interaktionsaufgaben unmittelbar einbezogen; in jeder x-beliebigen Interaktion, die über eine flüchtige Begegnung hinausgeht, sind Erwartungserwartungen in ganz unterschiedlichen Richtungen und Dimensionen im Spiel – je nach Aufgabe, die gerade im interaktiven Vordergrund steht. Fokussierungen lassen sich deshalb darauf beziehen, welche der immer gleichzeitig zu bearbeitenden Aufgaben gerade im Vorder- bzw. Hintergrund stehen. Daraus erwächst ein Interaktionsprofil abnehmender aufgabenorientierter Dringlichkeit, das die soziale Typik einer konkreten Interaktionsepisode ausmacht. Die soziale Typik ergibt sich kurz gesagt daraus, welche oder welcher der Aufgabenbereiche im Vorder- und welche oder welcher im Hintergrund stehen.

Das Zustandekommen eines Interaktionsprofils mit Vorder- und Hintergrund der Aufgabenerledigung hängt u.a. damit zusammen, wie die Ressourcen der Interaktion eingesetzt werden. Dass Interaktion sich über-

haupt so effizient und so ökonomisch als Problemlösungszusammenhang etablieren und überhaupt in dieser Ausprägung von Kopräsens möglich und wahrscheinlich werden kann, lässt sich nur vor dem Hintergrund der Ressourcen plausibel machen, auf die in und mit Interaktion zurückgegriffen werden kann. Wir konzentrieren uns an dieser Stelle, wie schon an anderer Stelle ausgeführt, auf Körper, Sprache und Architektur als drei grundlegende evolutionäre Errungenschaften. So kann man wissen – um nur drei plakative Beispiele illustrierend anzuführen –, dass die Ressourcenqualität der Sprache im Bereich der Syntax für das Problem des »turn taking« besonders ausgeprägt ist, die Ressourcenqualität des Körpers im Bereich der menschlichen Sinneswahrnehmung für die Probleme der Eröffnung und Situierung besonders ausgeprägt ist und etwa die Ressourcenqualität der Architektur im Bereich der gesellschaftlichen Funktionsräume für die Rahmung besonders ausgeprägt ist. Und man kann weiter wissen, dass der Einsatz von Sprache, pauschal gesagt, grundsätzlich einen Zugewinn an Dringlichkeit der jeweiligen Aufgabenerledigung mit sich bringt. Aufgabenerledigungen, die sich nur im Medium der Wahrnehmungswahrnehmung abspielen – wie z.B. die Wahrnehmung von Zugehörigkeit im Kontext der Positionierung –, sind grundsätzlich hintergründiger als solche, die eine sprachliche Form annehmen – wie z.B. das Anzeigen von Zugehörigkeit mittels der Verwendung deiktischer Ausdrücke. Und natürlich wäre dann auch noch zwischen verschiedenen sprachlichen Verfahren weiter zu differenzieren – nur im Falle der Klärung von Zugehörigkeit (Wo kommst Du her?), in der z.B. Herkunft zum Gegenstand einer eigenen Prädikation wird, erreicht die Positionierung ein Höchstmaß an Dringlichkeit und Vordergründigkeit.⁹¹ Das Beispiel zeigt, wie sich an dieser Stelle jeweils aufgabenspezifisch komplexe Forschungsfelder auftun, innerhalb derer das Verhältnis von Aufgaben- und Ressourcendifferenzierung empirisch zu beschreiben ist und in vielen Bereichen auch schon beschrieben worden ist. Wir werden darauf im Folgenden bei der Besprechung der einzelnen Aufgaben exemplarisch zurückkommen.

II.

Wir fangen mit dem Anfangen an. Beim Anfangen kann viel mehr schiefgehen (oder auch: *gar nicht* gehen), als man gemeinhin annimmt. Anfänge sind »zugleich notwendig und prekär«.⁹² Es heißt zwar anschaulich bei Luhmann: »Gesetzt den Fall, zwei oder mehr Personen geraten einander ins Feld wechselseitiger Wahrnehmung, dann führt allein diese Tatsache schon zwangsläufig

zur Systembildung.«⁹³ Aber wie genau »ger[ät] man einander ins Feld wechselseitiger Wahrnehmung«, wenn das nicht mehr oder weniger zwangsläufig (wie im Fahrstuhl) passiert, sondern in einer z.B. leicht unüberschaubaren Umgebung, in der es viele andere Reize gibt, die die Aufmerksamkeit auf sich ziehen? Wie können sich die Beteiligten überhaupt sicher sein, in ein Feld wechselseitiger Wahrnehmung gekommen zu sein? Wie wahrscheinlich ist es, dass sich ein Akt wechselseitiger Wahrnehmung in der freien Wildbahn ergibt, in der alle Beteiligten mit ihren eigenen Angelegenheiten beschäftigt sind (und z.B. mit guten Gründen Blickkontakt vermeiden)? Und selbst wenn Blickkontakt zustande kommt, ist doch nicht selbstverständlich, dass er ausreicht, ein genügend belastbares Feld wechselseitiger Wahrnehmung zu etablieren. Und wie ist sichergestellt, dass dieses Feld nicht nur etabliert, sondern auch aufrechterhalten wird und nicht gleich wieder zusammenbricht? Wie kann ich, mit anderen Worten, die Andere oder den Anderen dazu bringen, mir für eine ungewisse Zeit ihre oder seine Aufmerksamkeit zu schenken?

Schließlich kommt eine Fülle von Problemen ins Spiel, die damit zu tun haben, dass sich im Erfolgsfall der Etablierung wechselseitiger Wahrnehmung sofort weitere Fragen aufdrängen, die leicht als Schwellen der Entmutigung dazu beitragen könnten, sich besser erst gar nicht auf eine Interaktionseröffnung einzulassen. Z.B. kommt sofort die Frage auf, in was für eine Art von Austausch der entstandene Kontakt überführt werden soll und welche Form der Begrüßung dafür angemessen ist. Auch ist nicht sichergestellt, dass ein genügend großer Themenvorrat bereitsteht und angezapft werden kann. Bei all dem muss in der Regel auch bereits geklärt sein, ob sich die Beteiligten schon einmal begegnet sind, ob sie sich namentlich kennen und wie vertraut sie miteinander sind. Oft ist es mit Worten zudem nicht getan, so dass zwischen verschiedenen Formen körperlicher Annäherung entschieden werden muss, die vom distanzierten Stehenbleiben über das Nicken und den Gruß bis zum Handschlag, der Umarmung und dem Sichküssen gehen können. In allen Dimensionen drohen Fehlschläge und Gesichtsverluste, wenn Nähe und Distanz von einer Seite falsch eingeschätzt werden, eine Bekanntheit übersehen oder überschätzt wird. Man muss die im Falle des Anfangens typischerweise in Sekunden zu treffenden Entscheidungen nicht weiter ausmalen, um deutlich vor Augen zu haben (und vielleicht auch zu erinnern), was alles im Rahmen der Interaktionseröffnung auf eine nicht mehr rückgängig zu machende Art und Weise schiefgehen kann. Die Unwahrscheinlichkeit der Interaktion ist im Fall des Anfangs nicht so weit von unserer Alltagsintuition entfernt, wie das sonst typischerweise der Fall ist – was tatsächlich mit einer Besonderheit des

Anfangens zu tun hat: Sehr viele Fragen müssen in sehr kurzer Zeit (verdichtet) in und mit Interaktion beantwortet werden. In all den illustrierten Richtungen lassen sich ohne weiteres empirische Untersuchungen anschließen, um nachzuzeichnen, wie es angesichts dieser Entmutigungsschwellen immer wieder gewagt und erfolgreich in Angriff genommen werden kann, sich auf Interaktion einzulassen.

Die Reichhaltigkeit des diesbezüglichen empirischen Forschungsstands kann hier nicht aufgearbeitet werden.⁹⁴ Stattdessen sollen lediglich ein paar konkrete Teilprobleme der Interaktionseröffnung und etablierte Lösungsmöglichkeiten skizziert werden, um zu zeigen, wie dabei die verschiedenen Ressourcen der Interaktion ineinandergreifen und einander ergänzen. Bereits angedeutet haben wir das Problem der wechselseitigen Wahrnehmung, das im Grunde mit der Anforderung der Wahrnehmungswahrnehmung einhergeht. Die Errungenschaft des Blickkontakts, die evolutionär u.a. die Beweglichkeit des Augapfels voraussetzt, so dass die Blickrichtung von anderen gesehen und mehr oder weniger genau antizipiert werden kann, ist ein starkes Mittel, Wahrnehmungswahrnehmung zu suggerieren und Interaktion anlaufen zu lassen, was immer dann noch geschehen mag. Blickkontakt sorgt dafür, dass die Beteiligten nicht nur einander sehen, sondern sehen können, dass sie einander sehen.⁹⁵ Letztlich hat es wohl vor allem mit Blickkontakt zu tun, dass Personen einander so leicht, wie es oben hieß, »ins Feld wechselseitiger Wahrnehmung« geraten können – ohne dass damit aber suggeriert werden soll, dass Blickkontakt für die Interaktion(-seröffnung) unerlässlich sei. Allenfalls bringt der Blickkontakt die Reflexivität der Wahrnehmung (das Wahrnehmen des Wahrgenommenwerdens) besonders anschaulich, vielleicht sogar prototypisch zum Ausdruck. Schon das Telefongespräch zeigt, dass und wie die Unterstellbarkeit und Erwartbarkeit wechselseitiger Wahrnehmung auch auf andere Weise in genügend belastbarer Form sichergestellt werden kann. Und nur darauf kommt es an. Im Übrigen fällt auch der Blickkontakt nicht vom Himmel. Er ist nicht voraussetzungslos, sondern muss systematisch vorbereitet werden. Man kann ihn nicht erzwingen, aber auf verschiedene Weisen nahelegen. Dazu gehören vor allem Verfahren der Annäherung und Zuwendung (Proxemik),⁹⁶ die auf dem Einsatz des menschlichen Körpers als motorisch komplex navigierbarem Sensor beruhen und in spezifische Körperkonfigurationen von Kopräsenz einmünden (wie die sogenannte F-Formation).⁹⁷ Annäherung und Zuwendung können durch Interaktionsarchitekturen von Umgebungen der Begegnung auf vielfältige Weise unterstützt werden. Hierhin gehören die schon beschriebenen architektonischen Benutz-

barkeitshinweise, insbesondere die Navigationshinweise. Mit dem Ausnutzen körperlicher und architektonischer Ressourcen kann die Herstellung eines für Blickkontakt höchst förderlichen Milieus auf sehr effektive und zwingende Weise vorbereitet werden, um Interaktion über Wahrnehmungswahrnehmung anlaufen zu lassen. Die dafür prinzipiell notwendige Synchronisierung der beteiligten Wahrnehmungsprozesse kann in anspruchsvollen Fällen zudem technisch durch akustische Signale zusätzlich erleichtert werden. Der Gong, der im Hörsaal den Vorlesungsbeginn markiert, ist dafür ein gutes Beispiel. Aber auch das hörbar-vernehmbare Schließen der Tür kann eine solche Funktion erfüllen. Die Beispiele lassen sofort weitere Fragen nach anderen Situationen aufkommen. Sie illustrieren aber schon, worauf es ankommt: dass und wie beobachtbare Phänomene im Kontext von Interaktionseröffnungen tatsächlich als Antworten auf konkrete Fragen und Probleme rekonstruiert werden können. Das zeigt sich konkret auch, wenn man einen nur flüchtigen Blick auf die sprachlichen Phänomene der Eröffnung wirft und dabei auf Grüße und Begrüßungen stößt. Ohne hier in die Details zu gehen, finden sich an dieser Stelle nicht zufällig Paarsequenzen der Begrüßung (in unzähligen Variationen). Paarsequenzen, die sich durch eine besonders starke konditionelle Relevanz zwischen dem ersten und dem zweiten Element der Sequenz auszeichnen, sind als solche von entscheidender Relevanz, weil sie die Flüchtigkeit und Fragilität einer nur auf Wahrnehmungswahrnehmung beruhenden sozialen Situation auf einen Schlag in die Verlässlichkeit eines stabilen Austausches überführen. Mit ihnen erhält die Reflexivität der Wahrnehmung eine hörbare Gestalt. Zugleich tritt die Interaktionseröffnung als das relevante interaktive Problem in den Vordergrund: Es wird, wie oben beschrieben, zu der am dringlichsten zu bearbeitenden Anforderung, indem es in Form einer Paarsequenz mit der stärksten Form konditioneller Relevanz ausgestattet wird.

Paarsequenzen lösen also das Problem, Wahrnehmungswahrnehmung verlässlich zu etablieren und aufrechtzuerhalten und zugleich von Wahrnehmungsprozessen ein Stück weit abzulösen und unabhängig zu machen. Wer in Hörweite Anderer spricht, kann davon ausgehen, dass ihr oder sein Sprechen wahrgenommen wird, ohne sich dessen in der Regel immer wieder rückversichern zu müssen. In der Phase der Etablierung dieser Sicherheit dient die Paarigkeit der Begrüßung gewissermaßen dazu, sich diese Reziprozität der Wahrnehmung auf eine nicht bestreitbare Weise gegenseitig aufzuzeigen. An dieser Stelle kann man konkret sehen, wie die verschiedenen Ressourcen der Interaktion ineinandergreifen und wie die Sprache hier ihre besondere Funk-

tionalität ausspielen kann: Das Sprechen ist ein seinerseits so anspruchsvolles und unwahrscheinliches Ereignis von Formgebung, dass es unter Anwesenden nur sehr schwer überhaupt ignoriert werden kann – und selbst dann in der Regel als Interaktion verstanden werden muss (was die Unmöglichkeit von Selbstgesprächen unter Anwesenden begründet).⁹⁸ Gleichwohl gilt: Das Auftreten von Paarsequenzen setzt eine für Wahrnehmungswahrnehmung bereits hergerichtete soziale Situation voraus. Darin besteht die »späte Verbalität« innerhalb von Interaktionseröffnungen: »Sprache kommt erst zu einem relativ späten Zeitpunkt ins Spiel. Werden die Interaktionsbeteiligten verbal aktiv, kann man in der Regel analytisch nachweisen, dass sie sich [...] in der Regel bereits intensiv ausgetauscht und die Möglichkeiten für den verbalen Austausch aktiv hergestellt haben.«⁹⁹ Die Vorbereitung der (verbalen) Begrüßung ist dabei selbst schon Teil der Interaktion – und muss streng genommen als solche auch ihrerseits vorbereitet werden. Tendenziell entzieht sich der Beginn der Interaktion damit einer endlichen Fixierung: Die Intervalle werden immer kleiner, aber theoretisch auflösen lassen sie sich nicht. Eher greifen an dieser Stelle irgendwann Grenzen der Auflösung, was die Möglichkeiten der Beobachtung betrifft. Auch aus dieser Sichtweise heraus wird die Interaktionseröffnung zu einem höchst unwahrscheinlichen Unterfangen. Das staunenswerte Wunder liegt dann darin, dass und wie diese Unwahrscheinlichkeit immer wieder in die Sicherheit und Verlässlichkeit einer Interaktionsepisode überführt werden kann.¹⁰⁰

Wenn und in dem Ausmaß, in dem sich eine solche Interaktionsepisode nicht auf eine flüchtige Begegnung oder auf Prozesse von Ko-Orientierung und Ko-Ordination mit weitgehendem Wegfall weiterer Ko-Operation(en) beschränkt, zeichnet die Interaktionseröffnung ein oft beschriebenes hohes Maß an Verdichtung aus, das gerade Prozesse des Anfangens zu einem höchst beliebten Gegenstand der Interaktionsforschung macht, an denen man sehr viel über die fragliche Interaktionsepisode lernen kann. Verdichtung bezieht sich dabei auf die Kondensierung der Fallstruktur einer konkreten Begegnung, aber auch auf die starke Strukturiertheit der Interaktion: So »sind gerade der Anfang und das Ende eines Gesprächs [...] höchst sensible und komplexe interaktive Angelegenheiten. Wie so oft erweist sich gerade an scheinbar unbedeutenden Ereignissen, die schnell und nebenher erledigt werden, die hochgradige Strukturiertheit alltäglicher Interaktion.«¹⁰¹ Vor dem Hintergrund der Interaktion als einem »multitasking system«, in dem stets mehrere Aufgaben gleichzeitig bearbeitet werden müssen, lässt sich diese Strukturverdichtung der Interaktion an ihrem Beginn damit erklären,

dass im Grunde spätestens mit dem Übergang zur Begrüßung alle genannten Aufgaben und Probleme auf einen Schlag erledigt werden müssen. Man kann das insbesondere an der Aufgabe der Positionierung (darin eingeschlossen: der sozialen Kategorisierung) sehr anschaulich zeigen, weil schon die Realisierung des Grüßens hochgradig implikativ ist für das Verhältnis der Sichbegrüßenden, das gerade dabei ist, sich in einer Interaktionsepisode zu materialisieren. Das fängt mit der Frage des gegenseitigen (Wieder-)Erkennens an, geht über Fragen der Zugänglichkeit der oder des Anderen für einen längeren als nur flüchtigen Kontakt und reicht bis zu Fragen der sozialen Beziehung und des persönlichen Kontaktes zwischen den Beteiligten.¹⁰² Wie in einem Samenkorn die daraus aufgehende Pflanze scheint die interaktive Fallstruktur in der Gestaltung der Interaktionseröffnung vorgezeichnet.¹⁰³ Es ist deshalb kein Zufall, dass sich an dieser hochgradig implikativen Schwelle der Interaktionseröffnung in besonders hohem Maße Routinen entwickelt haben, die maßgeblich dabei helfen, den Anfang immer wieder zu wagen und diese Schwelle immer wieder zu überwinden. Die bereits herausgegriffenen Paarsequenzen sind deshalb nicht zufällig stark formelhaft ausgeprägt. Es sind die vorgefertigten Standardlösungen für ein offensichtlich wiederkehrendes kommunikatives Problem in all seinen Facetten, sprich: in der Komplexität der an dieser Stelle zu bearbeitenden Aufgaben.

Von Routinen ist es nur ein kleiner Schritt zu Ritualen, von der im Grunde nicht vermeidbaren Routinisierung der Interaktionseröffnung zu einer weitergehenden Ritualisierung, mit der die Interaktionseröffnung eine Aufwertung um ihrer selbst willen erfährt. Es geht also nicht nur darum, irgendwie anfangen zu müssen, indem die damit verbundenen Aufgaben (der Annäherung, der Vergewisserung wechselseitiger Wahrnehmung, der Synchronisierung, ...) gelöst werden. Es geht darum, das Anfangen auch *als* Anfang von etwas, *als* Beginn zu markieren. Je stärker das Anfangen die Form einer ausgebauten Eröffnungsphase annimmt, desto mehr wird die Aufmerksamkeit auf das Anfangen gelegt, das Anfangen selbst angefangen. Mit der Reflexivität des Anfangens geht nicht nur eine Aufwertung des Anfangs, sondern auch eine Aufwertung der folgenden Interaktionsepisode einher. Das Anfangen bekommt sozusagen immer mehr Eigenleben und verselbständigt sich mit markanter Ausgestaltung der Formen.

Verselbständigungen der Interaktionseröffnung zeigen sich auch dann sehr deutlich, wenn es – wie häufig in der institutionalisierten Kommunikation – zu performativen Äußerungen kommt, die das Begrüßen selbst auffällig hervortreten lassen: »Hiermit begrüße ich Sie sehr herzlich zu unserer heuti-

gen Vorlesung.«¹⁰⁴ Im Vordergrund steht dann nicht mehr die Dringlichkeit der Aufgabenerledigung (das konditionell relevant gemachte Zurückgrüßen), sondern schon abgelöst davon die Teilnahme an einer (feierlich ausgestaltbaren) Begrüßungszeremonie. Das ist hier mit Ritualisierung gemeint.¹⁰⁵ Mit Ritualisierung wird zwangsläufig Zeit verbraucht. Es dauert also länger, bis es mit dem losgeht, was auf die Eröffnung folgt. Darin ist kein Nachteil, sondern im Gegenteil ein gewünschter Effekt von Ritualisierung zu sehen. Roland Barthes hat in den »Mythen des Alltags« am Beispiel eines Auftritts von Billy Graham im Pariser Velodrome d'Hiver gezeigt, dass und wie der Aufschub und die Verzögerung dazu dienen, die fragliche Interaktionsepisode aufzuwerten. Je stärker die Eröffnung als Zeremonie zelebriert und feierlich in die Länge gezogen wird, desto stärker ist »in dieser ersten Phase der Zeremonie die große soziologische Bewegkraft der Erwartung« im Spiel.¹⁰⁶ Man kann auch sagen: Das Erleben von Kopräsens wird mit der Ritualisierung der Interaktionseröffnung strategisch aufgewertet. Es fängt nicht einfach an, sondern das Anfangen wird eigens markiert und hervorgehoben. Im Anschluss an eine Strategie aus der schriftlichen Kommunikation könnte man auch von einem *emphatischen* Interaktionsbeginn sprechen.¹⁰⁷ Die Interaktion hat schon längst begonnen, wenn der Beginn noch einmal emphatisch gefeiert wird. In diesen Fällen nimmt das Eröffnungsritual die Form einer Zeremonie an.

Auch unabhängig von solchen Fällen, die wie angedeutet häufig in der institutionalisierten Kommunikation anzutreffen sind, wo die zeremonielle Ausgestaltung der Begrüßung bereits zur Rahmung des Geschehens beiträgt, werden Interaktionseröffnungen auch sonst häufig zum Schauplatz mehr oder weniger alltäglicher Ritualisierungen, mit denen sich die Beteiligten nach Goffman ihrer wechselseitigen Wertschätzung und ihrer sozialen Beziehung versichern.¹⁰⁸ Hierhin gehören formelhafte Ausgestaltungen der Eröffnungsphase unterschiedlicher Art (Wie geht's?), insofern sie den Moment des Interaktionsbeginns gleichsam verlängern und ausdehnen. Wie häufig bei Ritualisierungen geht es dabei auch und gerade um Zugehörigkeit, die auf mehr abstellt als die auf Wahrnehmungswahrnehmung basierende soziale Anwesenheit, eben wahlweise auf etwas den flüchtigen Moment der Interaktion Überdauerndes wie soziale Gemeinschaft im Sinne von Bekanntheit, Vertrautheit oder Freundschaft innerhalb einer sozialen Beziehung. Kopräsensrituale im Rahmen der Interaktionseröffnung haben deshalb das Potential, die Grenzen von Kopräsens zu transzendieren, in diesem Fall: die Flüchtigkeit und zeitliche Begrenztheit der Interaktionsepisode zugunsten

einer schon früher andauernden und die Interaktion überdauernden sozialen Gemeinschaft. Genau das macht den Kern der Ritualisierung aus.¹⁰⁹

III.

Sobald eine Interaktionsepisode den Moment einer flüchtigen Begegnung überdauert und sich verfestigt, drängt mit Macht ein weiteres interaktionskonstitutives Problem in den Vordergrund: das Problem des *Sichabwechselns*, das bereits innerhalb der Eröffnungsphase im Spiel ist, wenn z.B. einer der Beteiligten grüßt und eine andere Beteiligte wenig später zurückgrüßt. Gleichwohl beruht die für Kopräsenz konstitutive Reflexivität der Wahrnehmung nicht auf einem Nacheinander aufeinander bezogener Wahrnehmungen, sondern auf der *Gleichzeitigkeit* von Wahrnehmungen, z.B. dem gleichzeitigen Sehen des Gesehenwerdens. Kopräsenz kann insbesondere im Rahmen von Ko-Orientierung und Ko-Ordination auch über die Eröffnungsphase hinaus und weitgehend auf multimodal realisierter Gleichzeitigkeit beruhen. Mit dem Übergang zur Ko-Operation im Rahmen gemeinsam geteilter Projekte (»joint projects«)¹¹⁰ wird Kopräsenz dann aber mehr und mehr ausgenutzt für anspruchsvolle, Zug-um-Zug-basierte Austauschprozesse. Das, was wir landläufig als Gespräch bezeichnen, ist ein solcher Zug-um-Zug-basierter Austauschprozess. Für ihn ist die Aufgabe des Sichabwechselns charakteristisch: Der oder die aktuell Sprechende muss sich darauf einstellen, zu dem oder der nächsten Hörenden zu werden, und umgekehrt muss sich der aktuelle Hörer, die aktuelle Hörerin darauf einstellen, zum nächsten Sprecher, zur nächsten Sprecherin zu werden. Das Sichabwechseln ist für den sprachlichen Austausch eine notwendige Folgelast: Das gesprochene Wort überdauert bekanntlich nicht den Moment seiner Hervorbringung und ist deshalb auf seine simultane Wahrnehmbarkeit (Hörbarkeit) angewiesen. Diese Wahrnehmbarkeit ist erheblich eingeschränkt, wenn nicht gleich ausgeschlossen, wenn mehrere Sprecher:innen gleichzeitig sprechen. Der Gebrauch von Sprache-in-Interaktion ist deshalb auf Reihenfolge und Nacheinander der Sprechenden angewiesen, anders gesagt: Sprecher:innen müssen sich abwechseln und ihr Nacheinander organisieren. Dazu sind sie nicht darauf angewiesen, die Reihenfolge ihrer Beiträge vorab zu klären und festzulegen (wiewohl auch das unter Anwesenden möglich ist). Sie müssen nicht einmal die Beiträge selbst vorab fabrizieren und dann nur noch abspulen (wiewohl auch das unter Anwesenden möglich ist). In sehr vielen Fällen gehen wir in eine Interaktion (»geraten einander ins Feld wechselseitiger Aufmerksamkeit«), ohne dass wir

uns vorher überlegt haben, was wir wann wie sagen wollen und wer wann »drankommen« soll. Selbstverständlich ist das freilich nicht. Im Gegenteil zeigt sich die Unwahrscheinlichkeit der verbalen Interaktion an dieser Stelle besonders anschaulich. Woher wissen die Beteiligten, wer als Nächste oder Nächster drankommt? Wie bringen sie es fertig, dass sie nicht immer wieder und über längere Zeit gleichzeitig sprechen? Wie vermeiden sie es umgekehrt, dass nicht immer wieder und über längere Zeit Schweigephasen entstehen? Woher können die Beteiligten wissen, wann die aktuelle Sprecherin, der aktuelle Sprecher »fertig« ist und ob eine Andere oder ein Anderer etwas sagen will? Wollte man Fragen dieser Art metasprachlich klären, käme der Austausch über Fragen der Regulation und Ordnung des Sprecherwechsels kaum hinaus. So geht es also nicht. Aber wie geht es dann?

Tatsächlich gibt es Interaktionen, in denen die aufgeworfenen Fragen weitgehend vorab geregelt werden und z.B. ein eigens für die Aufgabe der Rederechtszuweisung zuständige Funktion geschaffen und mit einer oder einem der Teilnehmenden besetzt wird, die oder der für nichts anderes als die Organisation des Sprecherwechsels (Moderation) zuständig ist, Listen führt und den »turn« jeweils übergibt an diejenigen, die ihr Redeinteresse auf eine vorab verabredete Weise »anmelden«. Es ist aber klar, dass wir es hier mit einer gleichsam domestizierten Form des Sichabwechsels zu tun haben, die nicht dem Alltag der freien Wildbahn entspricht (und im Übrigen auch nicht immer funktioniert). Näher an diesem Alltag ist die Einsicht, dass man das Sprechen-und-Zuhören nicht von außen an das Problem des Sichabwechsels herantragen muss (als käme es später dazu), sondern dass es von Anfang an – onto- wie phylogenetisch – in das Sichabwechseln unter Anwesenden eingebunden ist und als Kommunikationsmedium durch dieses Sichabwechseln maßgeblich geprägt worden ist. Dass sich mit dem Gespräch ein Zug-um-Zug-System des Sichabwechsels herausgebildet hat, ist also Teil der Evolution von (gesprochener) Sprache. Daher verwundert es nicht, dass die Sprache eine einzigartige Ressource für die Lösung der Teilprobleme des Sichabwechsels ist, mehr noch: für die Zwecke des Sichabwechsels »designt« worden ist. Sprechen unter Anwesenden impliziert deshalb eine fortlaufende Signalisierung nicht nur des aktuellen Dranseins, sondern auch der erwartbaren Dauer des weiteren Dranseins. Zuhören unter Anwesenden impliziert entsprechend eine fortlaufende Auswertung dieser Signale des Dranseins und der erwartbaren Dauer des Dranseins einschließlich insbesondere der Auswertung aller Signale für ein mögliches eigenes Drankommen. Wahrscheinlich ist es nicht übertrieben zu sagen: Sprechen-und-Zuhören *sind*

genau das, produzieren genau das: Signale für Dransein und Drankommen. Dabei gilt es, sich daran zu erinnern, dass unter Kopräsenzbedingungen Sendung und Empfang von Signalen nie einseitig verteilt sind: Jede Senderin und jeder Sender ist zugleich Empfänger:in, jede Empfängerin und jeder Empfänger zugleich Sender:in von Signalen.¹¹¹ Das Nacheinander von Sprechen- und -Zuhören, von Sprecher:in und Zuhörer:in lebt deshalb von der Gleichzeitigkeit der Multimodalität des Sendens und Empfangens von Signalen des Dranseins und möglichen Drankommens. Sprecher:innen liefern deshalb im Alltag keine vorgefertigt-vorfabrizierten Beiträge ab, sondern passen das, was »ihr Beitrag« werden wird, während der Produktion (»online«) an, so dass es im Kleist'schen Sinn tatsächlich zu einer »allmählichen Verfertigung der Gedanken beim Reden« kommt, zu einer »inkrementell« erfolgenden Produktion von Redeeinheiten.¹¹² Darin ist keine besonders anspruchsvolle (»künstlerische«) Leistung ausgewählter Sprecher:innen zu sehen, sondern ein unvermeidlicher Nebeneffekt der Produktion und Rezeption gesprochener Sprache unter Bedingungen von Kopräsenz: Sprecher:innen können nicht umhin (wenn sie nicht spezielle künstliche Vorkehrungen treffen) wahrzunehmen, wie das, was sie gerade von sich geben, von den Anwesenden auf- und wahrgenommen wird. Man kann sich vorstellen, dass es unter diesen Bedingungen höchst naheliegend ist, das eigene Sprechen entlang dieser Wahrnehmung zu modellieren. Das betrifft nicht nur, aber auch und gerade Signale, die sich auf das Sichabwechseln beziehen lassen, also z.B. auf das Drankommen-Wollen der Gesprächspartnerin oder des Gesprächspartners. Auf der anderen Seite können Zuhörer:innen ihrerseits nicht umhin, den im Werden befindlichen Beitrag des Sprechers oder der Sprecherin im Hinblick auf Signale des Dranbleibens auszuwerten, um das eigene Drankommen vorzubereiten und möglichst passend zu vollziehen, ohne die Andere oder den Anderen zu unterbrechen.

Wenn diese Beschreibungen halbwegs zutreffen, sollte sich insbesondere die Sprache in ihren grundlegenden Strukturen als Ressource des »turn taking« erweisen, die es z.B. auf Seiten der Zuhörerin oder des Zuhörers erlaubt, redeübergaberelevante Stellen nicht nur zu erkennen, sondern im Voraus antizipieren zu können, d.h. Erwartungen auszubilden, wie lange der oder die Andere noch dran sein wird und wann man selbst voraussichtlich dran kommen können wird, um diese Erwartungen durch entsprechende Signale zu etablieren und für die Sprecherin oder den Sprecher erwartbar zu machen. Auf Seiten der Sprecherin bzw. des Sprechers sollte es nicht nur möglich, sondern geradezu notwendig und unvermeidbar sein, mit dem Sprechen Signale des Dran-

seins und der Dauer des weiteren Dranseins auszusenden, also Erwartungen zu manifestieren und ihrerseits erwartbar zu machen, die das eigene weitere Sprechen, aber auch das Sprechen Anderer betreffen. Es ist nicht übertrieben zu behaupten, dass die Konversationsanalyse diese Nachweise in großer Fülle und Detailliertheit erbracht hat. Das gilt insbesondere für die Syntax der gesprochenen Sprache, die bereits in den ersten Arbeiten der Konversationsanalyse als wesentliche Ressource für die Konstruktion mehr oder weniger voraus-sagbarer und vorhersehbarer Redeeinheiten (»Turn«-Konstruktionseinheiten) (an)erkannt wurde. Syntax ist wohl nicht zuletzt und gerade das: ein System von Einheiten- und Abgrenzbarkeitshinweisen, das für Zug-um-Zug-Systeme wie das Gespräch konstitutiv ist – und wahrscheinlich genau deshalb entstanden ist:

Einer der Gründe, warum es in der menschlichen Sprache überhaupt so etwas wie Syntax gibt, scheint zu sein, dass die Syntax es ermöglicht, unabhängig vom Inhalt Vorhersagen über Abschlusspunkte von Redebeiträgen im Gespräch zu treffen. [...] Ohne Syntax und die damit verbundenen Projektionsmöglichkeiten wäre die erstaunliche Schnelligkeit und Reibungslosigkeit des Sprecherwechsels nicht erklärbar.¹¹³

Es versteht sich, dass die Syntax dabei nicht isoliert erscheint, sondern eingebettet in die gesamte Bandbreite der Multimodalität der Interaktion. Das betrifft auch weitere sprachbezogene Ressourcen wie die Prosodie, die Semantik und die Pragmatik von Äußerungen. Der Rückgriff auf diese unterschiedlichen Ressourcen kann von Fall zu Fall sehr unterschiedlich ausgeprägt sein. Es ist bekannt, dass in einer weitgehend an praktischen Tätigkeiten orientierten Ko-Operation die Sprache und alle sprachbezogenen Ressourcen stark zurücktreten können, während in gesprächslastigen (verbosen) Interaktionen, in denen die auf Sichtbarkeit beruhenden Modalitäten eingeschränkt oder wie am Telefon ganz ausgeschaltet sind, die Sprache als Ressource prägnant hervortritt. Wenn gesprochen-und-zugehört wird, ist grundsätzlich neben der Syntax auch die Prosodie als Ressource für Einheitenhinweise im Spiel. Vor allem Betonungen und Akzente, Tonhöhenbewegungen (insbesondere zum Einheitenende hin) und Sprechgeschwindigkeitsveränderungen tragen dazu bei, dass Intonationseinheiten entstehen, die als solche wahrgenommen und ausgewertet werden können für die Projektion einer Gesprächseinheit und ihres erwartbaren Abschlusses. Neben den formalen Einheitenhinweisen gibt es natürlich auch semantisch und pragmatisch definierte Einheitenhinweise, die

mit dem Bezug auf Welt und Handlung zu tun haben und ebenfalls erheblich dazu beitragen, dass das Sprechen-und-Zuhören tatsächlich der Erzeugung und Auswertung von Signalen des Dranseins und Drankommens entspricht. Es ist leicht vorstellbar, dass die Rekonstruktion von Einheitenhinweisen und ihres komplexen Zusammenspiels in der verbalen Interaktion ein weites Feld an Standardisierung und Variation eröffnet, das in einer mehr oder weniger situations- und kontextabstrahierenden, zudem oft stark schriftorientierten Grammatik – wenn überhaupt – nur ansatzweise erfasst werden kann. Darin ist einer der Gründe zu sehen, warum die Rezeption der Konversationsanalyse vor allem in der Linguistik auf einen sehr fruchtbaren Boden gefallen ist und mit der Gesprächsanalyse eine eigenständige linguistische Forschungsrichtung begründet hat, die nach wie vor auch und gerade mit Bezug auf das Sichabwechseln verfeinert und weiterentwickelt wird.¹¹⁴

In der Aufgabe des Sichabwechselns manifestiert sich auf besonders markante Weise ein allgemeines Strukturprinzip verbaler Interaktion, das Kopräsenz mit Nacheinander und Reihenfolge verknüpft (»Who comes next?«). Sprache dürfte als evolutionäre Errungenschaft maßgeblich dazu beigetragen haben, unter Bedingungen der auf Gleichzeitigkeit beruhenden Wahrnehmungswahrnehmung ein Zug-um-Zug-System zu etablieren, das weitgehend – wie am Sprecherwechsel soeben beschrieben – auf der Etablierung von Anschlussmöglichkeiten für nächste Züge beruht. Damit erhält das chronologische Vergehen von Zeit den Charakter sozial erlebter und behandelte Zeit, in der sich Erwartungserwartungen an das, was als Nächstes kommt, manifestieren können. In der Konversationsanalyse wird dieses Phänomen mit dem Terminus Sequentialität verbunden. Sequentialität meint viel mehr als nur Zeitlichkeit und Prozessualität.¹¹⁵ Über das Verstreichen von Zeit hinaus verweist Sequentialität auf eine grundlegend und genuin soziale Strukturebene der Selektivität von Kommunikation, die sich in der Zeit entfaltet. Kommunikation ist nicht nur ein zeitbindender Prozess, sondern ein Prozess, in und mit dem Zeit als Sinnfestlegung erfahrbar wird, die sich in der Erzeugung von Anschlussmöglichkeiten, -bedingungen und -erwartbarkeiten manifestiert. Im Sprecherwechsel kommt dieser Prozess wie die Spitze des Eisbergs an die Oberfläche.¹¹⁶

IV.

Der Aufgabe des Sichabwechselns steht die Aufgabe des Beitragens und der Themenorganisation nahe, dem »Wer-kommt-als-Nächste oder -Nächster?«

das »Was-kommt-als-Nächstes?«. Das Sichabwechseln der Sprecher:innen impliziert den Bezug (die Referenz) auf Welt und das Beitragen zu einem Thema, das das Gespräch zusammenhält und insbesondere in institutionalisierter Kommunikation auch den Anlass für die Begegnung darstellen kann. Bezüge auf Welt gehören zu Kopräsenz wohl von Anfang an dazu. Sie dürften zunächst aber kaum über das hinausgereicht haben, was für die Anwesenden situativ wahrnehmbar ist, vereinfacht gesagt: über das, worauf *gezeigt* werden kann und was auf diese Weise Teil der gemeinsam geteilten Aufmerksamkeit werden kann. Das ändert sich mit der Konfiguration von Kopräsenz durch Sprache, mit der es möglich wird, auf Welt auch außerhalb des sinnlich Wahrnehmbaren Bezug zu nehmen, Welt in einem umfassenden Sinn zu thematisieren. Ungeachtet dessen bleibt die Wahrnehmung – gerade zu Interaktionsbeginn – eine wichtige Ressource für den thematischen Einstieg in ein Gespräch und für das, worüber man relativ voraussetzungslos sprechen und wozu man beitragen kann.

Wie alle anderen Aufgaben schafft auch die Anforderung des Beitragens eigene Schwellen der Entmutigung, die zur Unwahrscheinlichkeit der Interaktion beitragen. Worüber sollen wir sprechen? Wie stimmen wir uns darüber ab, worüber wir reden wollen? Bekanntlich ist es dafür nicht notwendig, vorab eine Tagesordnung mit Besprechungspunkten festzulegen oder am Anfang einer Begegnung das Thema oder die Themen explizit auszuhandeln. Aber wie funktioniert es dann? Wie kommt man in ein Gesprächsthema hinein, und woher kann man wissen, worüber die oder der Andere gerne reden möchte und welches Thema auf keinen Fall angeschnitten werden sollte? Woher kann man wissen, dass mögliche Themen nicht schnell erschöpft sind und der gemeinsame Themenvorrat überhaupt genügend Stoff hergibt? Wer ist für die Themenwahl und -initiierung zuständig? Wieso führt die Frage, was als Nächstes kommen soll, nicht zu einem Dauerstress für die Beteiligten, der sie daran hindert, sich überhaupt auf ein Gespräch einzulassen? Wie beim Sichabwechseln beruht auch das Beitragen auf dem Aufbau und der Etablierung von Erwartungen, die in diesem Fall die Frage betreffen, *worüber* geredet werden soll. Es versteht sich, dass die Etablierung thematischer Erwartungen mit Erwartungen an die nächste Sprecherin oder den nächsten Sprecher einhergehen (können) und umgekehrt. Das Beitragen ist so gesehen Teil der Sequentialität der Interaktion, d.h. der Organisation von Reihenfolge und Nacheinander. Die thematische Struktur eines Gespräches ist deshalb grundsätzlich eine sequentielle Struktur (»why that now«),¹¹⁷ die sich in Themaeinführungs-, -beibehaltungs-, -entwicklungs-, -abschluss- und -wiedereinführungshinweisen

manifestiert,¹¹⁸ mit denen sich die Beteiligten aufzeigen, dass ein bestimmtes Thema auf eine bestimmte Weise an einer bestimmten Stelle des Gespräches eingeführt (»bounded topic«),¹¹⁹ beibehalten, weiterentwickelt (»shaded topic«),¹²⁰ abgeschlossen¹²¹ und womöglich noch einmal wiederaufgenommen wird. Worüber geredet wird, ergibt sich auf diese Weise in vielen Fällen alltäglicher Interaktion aus dem Gespräch selbst.

Wie genau das im Einzelnen funktioniert und welche Ressourcen dabei genutzt werden, ist in der Konversationsanalyse immer wieder untersucht worden.¹²² So ist bekannt, dass am Anfang der Interaktion der Bezug auf Welt weitgehend ritualisiert ist durch kürzere Austauschaktivitäten in Form von Paarsequenzen. Sie konstituieren eine Art »vorthematische Eröffnungssequenz«¹²³ im Sinne eines Vorlaufs, in dem es um allgemeine »interaktionsgeschichtliche Präliminarien«¹²⁴ geht, die noch nicht zu gesprächsspezifischen Themen gehören und deshalb den Themendruck zu Beginn des Gespräches erheblich entlasten und gleichzeitig auch schon Möglichkeiten für Expansionen bieten. Dazu können etwa die Vorstellung und Identifizierung gehören und der Austausch von Erkennungssignalen, insbesondere wenn es sich um Erstkontakte handelt, Erkundigungen nach dem Wohlbefinden und Ergehen der oder des Anderen und auch Fragen nach der Verfügbarkeit für ein Gespräch. Daneben macht sich in dieser Vor- und Frühphase des Beitragens auch die Wahrnehmung der Beteiligten als Ressource für ein thematisches »warming-up« bemerkbar. Unter bestimmten Umständen können Aspekte der sinnlichen Wahrnehmung, insbesondere wenn sie irgendwie auffällige und außergewöhnliche Umstände betreffen, leicht die Schwelle der Thematisierbarkeit überschreiten. Dazu gehört bekanntlich das Wetter, das sich auch unter Unbekannten als thematische Ressource in der Frühphase des Gespräches eignet.¹²⁵ Aufgrund ihres rituellen Austauschcharakters werden diese Sequenzen noch zur Gesprächseröffnung gezählt, wenngleich sich daraus, wie schon erwähnt, leicht erste thematische Expansionen ergeben können.¹²⁶ Den Übergang in die thematische Phase und die Aufgabe des Beitragens im engeren Sinn markiert dann die Einführung des »ersten Themas« (»first topic«).¹²⁷ Es ist nicht nur chronologisch das erste Thema, das im Anschluss an die Eröffnungssequenz mit ihren Austauschritualen gesprächsspezifisch angeschnitten wird. Es hat als solches auch einen besonderen Status als besonders wichtiges Thema, insofern es häufig den Anlass und Grund des Gespräches ausmacht und in dieser Funktion durch Gliederungssignale auch markiert wird. Mit der Initiierung des ersten Themas nimmt die Aufgabe des Beitragens erkennbar Fahrt auf. Aber natürlich braucht nicht jede Inter-

aktion einen Grund und Anlass, mit dem eine oder einer der Beteiligten in das Gespräch hineingeht. Kopräsens kommt auch zustande, indem sich die Beteiligten mehr oder weniger zufällig in das Feld wechselseitiger Wahrnehmung geraten und daraufhin ein Gespräch starten, das keinen vorgegebenen Zweck hat, sondern aus einer Gelegenheit heraus entsteht (Small Talk). Für die Anforderung, unter diesen Bedingungen geeignete Themen ausfindig zu machen, haben sich insbesondere in Erstkontakten unter einander nicht schon bekannt-vertrauten Gesprächspartnerinnen oder Gesprächspartnern sogenannte »which«-type questions« bewährt, die sich auf die kategorielle Zugehörigkeit der oder des Anderen beziehen. Zugehörigkeit ist eine Ressource für das Ausfindigmachen von Gesprächsthemen, wenn die Beteiligten nicht genügend voneinander wissen, um bereits über verlässliche und ergiebige Gesprächsthemen zu verfügen. In derartigen Situationen sind explizite Klärungen von Zugehörigkeit keine Seltenheit, weil Zugehörigkeiten in diesem Zusammenhang als Topoi (aufgrund) der Person fungieren, mit deren Hilfe konversationelle Anschlussmöglichkeiten für das Beitragen zu einem Thema ausfindig gemacht werden können. In diesem Zusammenhang berichtet Harvey Sacks über einen speziellen Typus von Fragen, die schon genannten »which«-type questions«, die die Klärung von Zugehörigkeit initiieren, indem sie als Antwort eine bestimmte soziale Kategorisierung verlangen, die die thematische Gesprächsfortsetzung erleichtern und ankurbeln: »When you get some category as an answer to a ›which‹-type question, you can readily formulate topics of conversation based on the knowledge stored in terms of that category.«¹²⁸ Die Klärung von Zugehörigkeit ist demzufolge häufig ein Motor der Themenprogression, der dazu beiträgt, Themenpotentiale aus einer Schritt-für-Schritt- bzw. »turn-by-turn«-spezifizierten sozialen Zugehörigkeit hervorgehen zu lassen und auf ihre Ergiebigkeit für das weitere Gespräch zu testen. Zugleich drängt mit der Klärung von Zugehörigkeit bereits die Aufgabe der Positionierung und sozialen Kategorisierung in den Vordergrund der Interaktion (auf die wir noch zurückkommen), mit der sich – allgemein gesprochen – die Gesellschaft als Themenressource bemerkbar macht, die dafür aufkommt, dass die Beteiligten ihre thematischen Ressourcen nicht ad hoc aufbauen und entwickeln müssen. Die Gesellschaft stellt so gesehen einen Themenvorrat zur Verfügung, der die Vielfalt der möglichen Themen einschränkt mit Bezug auf besonders naheliegende und erwartbare Themen. Niklas Luhmann spricht in diesem Sinne von »Kultur« und »Semantik«:

Themen werden nicht jeweils fallweise neu geschaffen, sind aber andererseits auch nicht durch die Sprache, etwa als Wortschatz, in ausreichender Prägnanz vorgegeben (denn die Sprache behandelt alle Worte gleich und disponiert noch nicht über die Themafähigkeit in kommunikativen Prozessen). Es wird demnach ein dazwischenliegendes, Interaktion und Sprache vermittelndes Erfordernis geben – eine Art Vorrat möglicher Themen, die für rasche und rasch verständliche Aufnahme in konkreten kommunikativen Prozessen bereitstehen. Wir nennen diesen Themenvorrat Kultur und, wenn er eigens für Kommunikationszwecke aufbewahrt wird, Semantik.¹²⁹

Kultur und Semantik gehören zu den wichtigsten Ressourcen für das Beitragen der Teilnehmer:innen zu Gesprächsthemen. Sie gehen über die grammatisch-syntaktischen Ressourcen der Sprache hinaus, mit denen z.B. thematisch relevante Referenzen im Zuge der Themaeführung auffällig hervorgehoben werden können,¹³⁰ indem sie erklärbar machen, dass und wie sich das Finden und Bearbeiten von Themen im Gespräch scheinbar wie von selbst ergibt. So versteht sich, dass die Beteiligten nicht davor zurückscheuen, sich auf die Anforderung des Beitrags immer wieder und auch ohne thematischen Vorsatz auch unter Unbekannten einzulassen. Zugleich wird im sprachlich konstituierten gesellschaftlichen Themenvorrat und seiner Nutzung beim Beitragen manifest, dass Interaktion »kein Staat im Staate«¹³¹ ist, sondern Mitvollzug von Gesellschaft: Auch unter Bedingungen von Kopräsenz sind die, die sozial anwesend sind, zugleich in komplexer Weise zugehörig zu sozialen (Wir-)Gruppen, die Anwesenheit und episodisch gebundene Kopräsenz überdauern, worauf wir bei der Besprechung der Aufgabe des Positionierens im Folgenden näher eingehen.

V.

Mit den »which-type questions« haben wir bereits die Aufgabe des *Positionierens* angesprochen. Die mit diesen Fragen verbundene Klärung von Zugehörigkeit innerhalb einer Frage-Antwort-Sequenz stellt freilich bereits eine sehr auffällige Umgangsweise mit der ansonsten oft im Hintergrund miterledigten Aufgabe der Selbst- und Fremddarstellung dar. Die Anforderung, sich selbst und die Anderen darzustellen und zu präsentieren, dürfte unmittelbar aus der Situation der Kopräsenz mit wechselseitiger Wahrnehmung hervorgegangen sein. In dem Maße, in dem die Beteiligten füreinander wahrnehmbar und aufeinander aufmerksam werden, machen sie sich auch in ihrer Präsenz zugäng-

lich und mit Bezug auf die oder den jeweils Anderen empfänglich für eigene und fremde Aspekte ihrer Identität und der ihrer Partner:innen, die über die wechselseitige Wahrnehmung bloßer Anwesenheit hinausgehen. In einem allgemeinen Sinn ist das Positionieren deshalb eine allgegenwärtige Praxis, die geradezu unvermeidbar anläuft und dabei auf Klärungsfragen wie die eingangs genannten in der Regel nicht angewiesen ist. Vor allem in der Soziologie und Sozialpsychologie ist das gut bekannt, wie die beiden folgenden Zitate belegen: »Wenn ein Fremder uns vor Augen tritt, dürfte uns der erste Augenblick befähigen, seine Kategorie und seine Eigenschaften, seine ›soziale Identität‹ zu antizipieren.«¹³² Und von der Aussage ähnlich: »We only have to look as far as the first street corner, where long hair, a police uniform or skin colour warrant quick assignment to a category from which further inferences can then be drawn.«¹³³ Kategorisierungen dieser Art, die wir zum Positionieren zählen (s.u.), sind offenbar alltäglich und unvermeidlich. Sie sind folgenreich und wirksam, insofern sie direkt zur Gruppenbildung beitragen, die als tiefe gesellschaftliche Implikation der Aufgabe des Positionierens zu verstehen ist. Auch das ist gut bekannt: »Die Bildung einer menschlichen Gruppe ist [...] ein symbolischer, nicht ein physischer Sachverhalt.«¹³⁴

Insbesondere mit Blick auf die Konstitution eines Wir, dem Andere (ihr) gegenübergestellt und von dem Andere abgegrenzt werden können (sie), ist die Kraft alltäglicher Kategorisierungen oft beschrieben worden. Kategorisierungen können sich dabei weitgehend freimachen von »natürlichen« Beschränkungen:

[D]ie Wir-Gruppen-Konstruktion [wird] als ein extrem variabler, ohne Muß- und Idealgrößen territorialer oder zahlenmäßiger Art auskommender Vorgang in der Reihe moderner ›Fabrikationen der Fiktionen‹ (Carl Einstein) ausgedeutet – als Erfindung einer Tradition, als Sinngebung absolut kontingenter Existenz, als Erzählung. [...] Es kann sich dabei um fiktive Abstammungsgemeinschaften und erweiterte Verwandtschaftsgruppen handeln, aber ebenso kommen Alters- und Generationskohorten, Religionsgemeinschaften, korporative Verbände und räumliche Nachbarschaften in Frage. Über alle solchermaßen konstruierten Wir-Gruppen werden Ressourcen und Besitzverhältnisse bestimmt, konstituiert sich Teilhabe an der Macht und werden moralische Instanzen organisiert. Ein selbstreferentielles, für selbstverständlich genommenes Zugehörigkeitsgefühl, das Rechte schafft, Pflichten begründet und vor allem Vertrauen schafft, kann sich einer endlosen Palette von inputs und symbolischer Formen bedienen. Die Wir-Gruppen-Konstruktion ist demnach ein elementares, zweiwertiges Klassifi-

kationsmuster. Der Code »wir und sie« (oder innen und außen) funktioniert ähnlich exklusiv wie die Geschlechterdifferenz, die politische Rechts-Links-Dichotomie, die Hierarchie von oben und unten oder der Gegensatz von Zentrum und Peripherie.¹³⁵

Natürlich vollziehen sich solche Prozesse der Gruppenbildung nicht ausschließlich und in der Gegenwartsgesellschaft vielleicht nicht einmal mehr primär unter Bedingungen von Kopräsenz. Aber wenn man sich fragt, wo und wie für solche voraussetzungsreichen Gruppenbildungen phylo- und ontogenetisch die Grundlagen gelegt werden, stößt man auf die Kernaufgabe des Positionierens, die als Anforderung bereits aus frühen Formen von Ko-Operation hervorgegangen sein dürfte. Dabei darf man sich den Sachverhalt allerdings nicht schon durch die Brille der Gegenwart(-sgesellschaft) heraus vorstellen. Es ist nicht so, dass die Beteiligten eben einfach vieles von ihrem Gegenüber sehen, was mit seiner und ihrer – als mehr oder weniger gegeben – gedachten »Identität« zu tun hat und womöglich darüber Aufschluss gibt, was für eine oder einer die oder der Andere und man selbst jeweils »ist« (z.B. gemäß uns Gegenwartsmenschen gewohnten sozialen Kategoriensammlungen wie z.B. Geschlecht, Alter, Herkunft, Konfession, Beruf oder Klasse). Vielmehr sind solche Rückschlüsse und Interpretationen auf der Grundlage von Zugehörigkeitsdarstellungen¹³⁶ bereits das voraussetzungsreiche Ergebnis von Kopräsenzsituationen, innerhalb derer die Beteiligten allmählich damit begonnen haben dürften, auf die für ihre jeweilige Kooperation relevanten Seiten der oder des Anderen (und an sich selbst) aufmerksam zu werden. Die Reichhaltigkeit der sozialen Kategorien geht mit der Komplexität der Ko-Operation Hand in Hand. Je komplexer die kopräsent erlebte und vollzogene Ko-Operation, desto reichhaltiger und vielfältiger die Seiten der oder des Anderen und die an sich selbst erfahrenen Seiten, die bei neuerlichen Zusammenkünften beobachtet und mehr und mehr auch erwartet werden können. Es ist anzunehmen, dass routinisiert eingespielte Ko-Operationen im Format gemeinsamer Aktivitäten dann auch allmählich ihre Spuren in Selbst- und Fremdbildern hinterlassen, die ihrerseits Formen annehmen können: Formen des ritualisierten Umgangs miteinander, Formen also des Verhaltens, aber auch Formen der Präsentation, in denen man sich der oder dem Anderen in einem bestimmten Licht zeigt und Aktivitäten in einem auf die Andere oder den Anderen angepassten Zuschnitt ausführt. Man zeigt sich (vor dem und für die Andere oder den Anderen, aber auch für sich selbst) mit dem, was man tut, und vor allem auch damit, *wie* man es

tut. All das hinterlässt Spuren im Auftreten, in Gebärden und in der Art, wie man sich hinstellt (eine Position einnimmt), und vielleicht auch in der Art, wie man seine äußere Erscheinung (Kleidung, Schmuck, Kosmetik, ...) auf eine bestimmte Kopräsenssituation hin vorbereitet und zurechtmacht. All das wiederum geschieht nicht in Isolation, sondern als Bestandteil geteilter Aufmerksamkeit. Leicht kann man sich vorstellen, wie im Kontext einer auf einen bestimmten Zweck zugeschnittenen Ko-Operation aus geteilter Aufmerksamkeit (für ein Projekt) so etwas wie »geteilte Intentionalität« wird mit Aufgaben, Zuständigkeiten und Verantwortungen, aus einem Wir, das hier und jetzt den Augenblick teilt, ein kollektives Wir, das über den Augenblick hinaus Werte und Normen bei »uns« unterstellt und voraussetzt – womit wir dann bereits bei frühen Formen von Wir-Gruppen-Bildung wären.¹³⁷ An dieser Stelle zeigt sich anschaulich, wie das Problem der Selbst- und Fremddarstellung (Wer sind wir?) und die darauf bezogene Aufgabe der Positionierung unmittelbar aus dem Problem der Situierung und der Aufgabe der Ko-Operation hervorgehen. Als direkte Folge von Ko-Operation trägt das Positionieren jedenfalls maßgeblich zu Prozessen der gesellschaftlichen Gruppenbildung bei und hilft erklären, wie aus anwesenheitsgebundenen Kopräsensgemeinschaften Gruppen entstehen können, die die Zufälle und Unwägbarkeiten von Kopräsensepisoden im Sinne sozialer Gruppen überdauern, die nicht auf Kopräsens, sondern auf Gesellschaft verweisen. Wie schon angedeutet bei der Besprechung der Interaktionseröffnung (s.o. Abschnitt II. in diesem Kapitel) ist u.a. das Positionieren dafür verantwortlich, dass sich Interaktion als Vollzug von Gesellschaft verwirklichen kann.¹³⁸ Je stärker die Entwicklung von Interaktion, Organisation und Gesellschaft und mit ihnen die Eigenständigkeit der damit verbundenen Bildung sozialer Systeme vorangeschritten ist, desto anspruchsvoller, unüberschaubarer, vielfältiger und reichhaltiger werden die Positionierungsmöglichkeiten und -herausforderungen, die in die Interaktion eingeschwemmt werden. Man sieht daran sehr anschaulich, wie sich die Interaktion mit dem Positionieren Potentialen von Zugehörigkeit öffnet, die jenseits von Anwesenheit liegen und auf Gesellschaft verweisen. Das zeigt sich auch schon an den sozialen Kategorien, die in den sogenannten »which-type questions« von den Teilnehmenden genutzt werden können, um Themen zu generieren für die Gesprächsfortsetzung. Die fraglichen sozialen Kategorien (die wir eingangs schon zitiert haben: Alter, Geschlecht, Herkunft, Konfession, Beruf, Klasse, ...) können ihre Rolle als Themenauslöserinnen nur spielen, weil sie als bekannt und relevant mehr oder weniger vorausgesetzt werden können (was nicht heißt, dass man mit ihnen nicht Schiffbruch

erleiden kann)¹³⁹ und in einer bestimmten Gesellschaft als »Kategoriensammlung« und «-inventar»¹⁴⁰ weit verbreitet sind.¹⁴¹ Gesellschaften zeichnen sich womöglich genau dadurch aus, was für Sammlungen und Inventare sie für das Positionieren, Kategorisieren, Typisieren, Gegenüberstellen, Zuordnen, Zuschreiben und Bewerten von Mitgliedschaften bevorzugen und vorrangig zur Verfügung stellen, kurz gesagt: nach welchen Kriterien und Maßstäben sie die soziale Welt aufteilen.¹⁴²

Die Breite der Aufgabe des Positionierens, die wir hier angefangen von ersten Sedimentierungen früher Zweck- und Intentionsgemeinschaften bis zur Komplexität sozialer Kategorisierungen unserer modernen Gesellschaft skizziert haben, spiegelt sich in der Vielfalt der Terminologie zur Bezeichnung und Untersuchung der Phänomene, die der Sache nach zu dieser Aufgabe gehören – wie die Reihung der Verben im letzten Satz des vorausgehenden Absatzes bereits andeutet. Wir halten hier am Positionierungsbegriff fest, weil es sich um einen vergleichsweise dehnbaren und zugleich anschaulichen Begriff handelt. Anders als die Interaktionseröffnung, der Sprecherwechsel und die Themenorganisation ist das Positionieren keine konversationelle Aufgabe, die in den Einführungen in die Konversationsanalyse wie selbstverständlich als eigene Aufgabe mitbesprochen wird. Dabei hat diese Aufgabe in der frühen Konversationsanalyse mit Harvey Sacks Studien zu den sogenannten »membership categorization devices« eine lange zurückreichende Forschungstradition, die in der Rezeption vor allem mit dem Konzept der sozialen Kategorisierung verbunden wird. Positionieren erscheint dann als eine Aufgabe, bei der es vor allem um das Zuordnen von Personen zu sozialen Gruppen und um das Zuschreiben und Bewerten gruppenspezifischer Eigenschaften und Verhaltensweisen geht.¹⁴³ Der Positionierungsbegriff kommt dagegen aus einer sozialpsychologisch geprägten Forschungstradition.¹⁴⁴ Dort erscheint das Positionieren eher im konkret-anschaulichen Sinn des Wortes als Art und Weise, sich gegenüber anderen im Gespräch zu »positionieren«, d.h. eine bestimmte Haltung und Position (z.B. für oder gegen eine Einstellung oder Haltung) einzunehmen. Dieser Positionierungsbegriff ist auch in nicht metaphorischer Weise darauf bezogen worden, wie Interaktionsbeteiligte in einem Raum bestimmte räumliche Positionen beziehen, sich dazwischen hin und her bewegen, andere verdrängen bzw. auf Plätze verweisen und selbst Plätze einnehmen.¹⁴⁵ Schließlich wäre auch noch auf die gesprächsanalytisch viel rezipierte Begrifflichkeit der Selbst- und Fremddarstellung hinzuweisen, die auf Erving Goffmans Konzept der Selbstdarstellung im Alltag zurückgeht¹⁴⁶ und z.B. bei der Analyse »[n]ationale[r] Selbst- und Fremdbilder« im Zuge der Umbruchsitua-

tionen in Ostmitteleuropa in den 1980er und 90er Jahren oft wiederaufgenommen worden ist.¹⁴⁷ Als Selbst- und Fremddarstellung erscheint Positionierung als eine Art unvermeidliche Begleiterscheinung des Sprechens-und-Zuhörens, die nicht nur aus dem resultiert, *was wir sagen*, sondern vor allem daraus, *wie wir es sagen*, und die maßgeblich an der Beziehungsdefinition der Teilnehmer:innen beteiligt ist.¹⁴⁸ Positionieren lebt als Aufgabe in diesem Zusammenhang vor allem davon, dass sie sich eher implizit und zwischen den Zeilen ergibt, also interaktiv hintergründig erledigt wird. Das muss natürlich nicht der Fall sein. Wie jede Aufgabe kann auch die Positionierung in den Vordergrund der Kommunikation geraten. Das ist vor allem bei Formen der Selbstthematization (und Fremdthematization) der Fall, in denen das eigene Selbst ausdrücklich zum Thema gemacht wird.¹⁴⁹

Der Durchgang durch die Breite der mit Positionierung verbindbaren Konzepte ist keineswegs abgeschlossen,¹⁵⁰ aber er dürfte andeuten, wie komplex und vielschichtig diese Aufgabe ist. Vor der Folie dieser Aufgabe entfaltet sich eine spezifische Unwahrscheinlichkeit der Interaktion daraus, dass die Beteiligten zu gemeinsam Beteiligten an einer Aufführung werden, die für alle Anwesenden verschiedene Positionen vorsieht, die von diesen gleich mit Beginn der Interaktion einzunehmen sind. Je komplexer die Gesellschaft ist, in der sich Kopräsens ereignet, desto vielfältiger und unvorhersehbarer werden die möglichen Präsenz- und Beteiligungsfiguren, die – zumal wechselseitig – in Rechnung zu stellen und zu übernehmen sind. Woher weiß ich, von welcher konkreten Seite ihrer oder seiner Identität sich die oder der Andere nähert und ich selbst mich ihr oder ihm nähern soll/will? Worauf muss ich in der Interaktion mit ihr oder ihm Rücksicht nehmen und welche Aspekte unserer Beziehung gilt es, in der Begegnung in den Vordergrund zu stellen? Was für eine Art von Gruppenbildung werde ich mit ihr oder ihm eingehen können? Als was präsentiere ich mich selbst und welche der vielen Aspekte unserer Zugehörigkeit(en) sollen in der Episode zum Tragen kommen? Welche Haltungen und Ansprüche soll ich der oder dem Anderen gegenüber einnehmen? Wer bin ich – und wenn ja, wie viele? Fragen wie diese, die in einer identitätssensiblen Gegenwartsgesellschaft vielleicht gar nicht so kontraintuitiv klingen (wie sie gemeint sind), deuten positionierungstypische Schwellen der Entmutigung an, die danach fragen lassen, warum und aufgrund welcher Ressourcen solche Schwellen im Alltag immer wieder überwunden werden.

Wenn wir uns bei der Kommentierung der Ressourcen an empirischen Studien aus unserer historisch rezenten Gegenwart orientieren, ist es naheliegend, zunächst auch in diesem Zusammenhang auf Wahrnehmung als

Ressource hinzuweisen. Die Wahrnehmung von Zugehörigkeit erlaubt eine vergleichsweise niedrigschwellige und hintergründige Miterledigung des Positionierens. Die Beteiligten sehen und hören sich – und haben mit ihrer wechselseitigen Wahrnehmung bereits weitreichende Zugehörigkeitsschlüsse gezogen, auf deren Grundlage sie die Interaktion in einer Weise fortsetzen können, dass darin basale Kategorisierungen bereits impliziert sind. Man denke daran, wie sehr eine bestimmte Situation, die äußere Erscheinungsform und Zurechtmachung des Körpers und weiterreichende Kleidungs- und Gestaltungsmöglichkeiten Zugehörigkeiten hochgradig nahelegen können, ohne dass darüber gesprochen werden muss. Freilich bleiben Positionierungen, die lediglich auf der Erwartbarkeit von Wahrnehmungen beruhen, typischerweise vage und unbelastbar. An dieser Stelle schafft die natürliche Sprache schon mit der Grammatik der Personalreferenz und dann noch einmal mit der weitergehenden Möglichkeit der expliziten Thematisierung Klarheit.¹⁵¹

Aber wie immer verändern die Aufgaben ihre Charakteristik und ihr Aussehen, wenn und sobald Sprache bei ihrer Erledigung die Hauptrolle übernimmt. Das gilt auch und gerade für das Positionieren, wenn man daran denkt, welcher Schub in die Selbst- und Fremddarstellung und in die damit verbundenen Prozesse der Gruppenbildung hineinkommt, wenn Zugehörigkeiten nicht länger mit augenblicks- und ortsbegrenzten Kollektivitäten verbunden werden, sondern auf Begriffe gebracht werden, d.h. über das »Wir« (und »Sie«) hinausgehen zugunsten der Semantik zugehörigkeitsbezogener Begrifflichkeiten. An historischen Kategorien zur Unterscheidung und Gegenüberstellung von Gruppen lässt sich diese Dynamik besonders gut studieren. So hat der Historiker Reinhart Koselleck speziell die Kraft von Begriffen als Indikatoren und Faktoren der Gruppenbildung herausgestellt. Begriffe gewinnen demnach ihre Eigentümlichkeit daraus, wie sie einerseits auf die Erfahrung mit »politisch wirksamen Handlungseinheiten« reagieren und wie sie andererseits die Erwartung an solche Handlungseinheiten auf einen Begriff bringen, der sich der zeitgenössischen Erfahrung zunächst gerade noch entzieht und deshalb zwangsläufig vorgreift.¹⁵² Koselleck weist entsprechend darauf hin, dass für die Bildung von sozialen Gruppen im Sinne politisch wirksamer Handlungseinheiten »Begriffe« notwendig sind,

in denen sich eine Gruppe wiedererkennen und selbst bestimmen muss, wenn sie als Handlungseinheit will auftreten können. Ein Begriff in diesem hier verwendeten Sinne indiziert nicht nur Handlungseinheiten, er

prägt und schafft sie auch. Er ist nicht nur Indikator, sondern auch Faktor politischer oder sozialer Gruppen.¹⁵³

Im Gegensatz zum »schlichten Gebrauch des ›wir‹ und des ›ihr‹«¹⁵⁴ zeichnet es Begriffe – das sind in Bezug auf Zugehörigkeit soziale Kategorien – aus, dass eine politische oder soziale Handlungseinheit damit »Kollektivnamen wie ›Nation‹, ›Klasse‹, ›Freundschaft‹, ›Kirche‹ usw.« erhält, die jeweils einen »konkreten Allgemeinheitsanspruch« zum Ausdruck bringen.¹⁵⁵ Erst mithilfe solcher Begriffe werden »Aus- und Eingrenzungen« auch über den Kreis der gerade Anwesenden, über das Hier und Jetzt von »wir« und »ihr« möglich und wahrscheinlich. Anders gesagt: »Die Wirksamkeit gegenseitiger Zuordnungen steigert sich geschichtlich, sobald sie auf Gruppen bezogen werden«.¹⁵⁶

Sie steigert sich noch einmal erheblich, sobald diese gruppenbezogenen Zuordnungen ihrerseits auf Kollektivnamen bezogen werden, in denen sich gruppenkonstitutive Einheit auch »begrifflich artikuliert«.¹⁵⁷ Soziale Kategorien dienen also nicht nur der Bezeichnung und Benennung der Gliederung der sozialen Welt, sondern genauso ihrer Bildung und Ausprägung. Die Verwendung und Schaffung von Selbstbezeichnungen im Gespräch, mithilfe derer sich Anwesende als Zugehörige einer sozialen Gruppe als solche wiederfinden, erleben und engagieren können, ist insofern ein Beleg dafür, wie auch unter den beschränkten Bedingungen von Kopräsens dazu beigetragen wird, politische bzw. soziale Handlungseinheiten zu schaffen und auszufüllen.

VI.

Die Aufgabe der *Rahmung* hängt eng mit der Aufgabe der Positionierung zusammen, geht aber nicht darin auf. Mit der Aufgabe der Rahmung ist ergänzend zur Positionierung gemeint, dass sich die Anwesenden Gewissheit darüber verschaffen, was, salopp gesagt, zwischen ihnen gerade abläuft (»What's going on?«). *Wer* sie gerade füreinander sind, hängt eng mit dem zusammen, *was* sie gerade zusammen miteinander machen (und umgekehrt). Insofern gehen Positionierung und Rahmung in der Regel Hand in Hand. Die Rahmung beinhaltet z.B. schon die Frage, ob es sich um fokussierte Interaktion handelt, bei der die Beteiligten ein gemeinsames Projekt mit wechselseitigem Austausch verfolgen, oder um unfokussierte Interaktion, bei der sich die Kooperation weitgehend auf Ko-Orientierung und Ko-Ordination beschränkt (wie z.B. in einer Wartezimmersituation). Natürlich steigen die Rahmungsanforderungen in fokussierter Interaktion noch einmal an, z.B.

mit Bezug auf die Frage, ob es einen speziellen Anlass und Grund des Treffens (»social occasion«)¹⁵⁸ gibt, wie im Falle institutionalisierter Interaktion (und worin der gegebenenfalls besteht), oder ob es sich um eine mehr oder wenige offene, z.B. dem Small Talk gewidmete Begegnung handelt, bei der der Ausgang noch offen ist. Rahmung findet in der institutionalisierten Interaktion einen besonders markanten Ausdruck (worauf wir noch zurückkommen werden), aber sie beschränkt sich nicht darauf. Wenn wir den Situationsbegriff nicht für die Aufgabe der Situierung reserviert hätten (s.o.), könnte man auch sagen, dass die Rahmung auf eine Art von Situationsdefinition hinausläuft: Die Beteiligten müssen sich im Klaren darüber sein, in was für einer sozialen Situation sie sich gerade befinden – und davon ausgehen können, dass sie diese Definition der sozialen Situation bis zu einem bestimmten Grad teilen. Das fängt eben nicht erst bei institutionell konfigurierter Kopräsenz an, sondern mit einer Art zumeist stillschweigender Begründung (»account«) für das aktuelle Zusammensein. Diese Begründung wird nicht ein für alle Male »gegeben«, sondern entsteht, wächst, entwickelt und verändert sich flexibel und dynamisch mit der Interaktion mit. Wenn man auch in diesem Punkt an frühe Formen von Rahmung herankommen will, darf man sich dazu nicht die in unzählige Aktivitätstypen sehr fein ausdifferenzierte moderne Gesellschaft vorstellen. Vielmehr wird man an einen noch gut überschaubaren gesellschaftlichen Zusammenhang ausgewählter gemeinschaftlicher Veranstaltungen denken (in enger Anlehnung an die Aktivitäten der täglichen Lebensnotwendigkeiten), die dem Beisammensein seinen sozialen Sinn, d.h. seinen für alle mehr oder weniger einsichtigen gemeinsamen Handlungsrahmen (»frame«) und den dazugehörigen Handlungsablauf (»script«) vermittelt haben. Die Aufgabe der Rahmung besteht dann eher in der Form einer unauffällig-unaufwendigen Kontextualisierung dessen, was gerade geschieht. Mitgebrachte und auch von Anderen gesehene Gegenstände, aber auch die Orte und Plätze, an denen man zusammenkommt, können im Sinne eines »Kontextualisierungshinweises«¹⁵⁹ reichen, die Interaktion als einen bestimmten Aktivitätstyp hinreichend festzulegen. Wahrscheinlich muss man davon ausgehen, dass sich relativ schnell mit der Wiederholung von bestimmten Formen der Kooperation auch entsprechende Interaktionsmuster herausgebildet haben, um die Frage, was gerade läuft, hinreichend zu beantworten und zur Grundlage voraussetzungsreicherer Aktivitätstypen zu machen.

Die Unwahrscheinlichkeit der Interaktion, die aus der Anforderung der Rahmung erwächst, hat mit dem notwendig sozialen Charakter von Koprä-

senz zu tun. Woher wissen die Beteiligten, in was für einen Aktivitätstyp sie gerade hineingeraten sind und dass ihre Auffassungen von dem, was abläuft, übereinstimmen? Sind alle Beteiligten für die Situationsdefinition gleichermaßen zuständig oder übernimmt eine oder einer der Beteiligten eine Führungsrolle? Wie zeigen sich die Beteiligten den gerade relevanten Gesprächsrahmen auf und wie kommen sie im Zweifelsfall von einem bestehenden zu einem neuen und anderen Gesprächsrahmen? Mit welchen und wie vielen Aktivitätstypen muss ich rechnen und wie bereite ich mich darauf vor? Fragen wie diese könnten leicht dazu führen, sich auf das soziale Wagnis einer Interaktion gar nicht erst einzulassen. Offensichtlich ist es nicht nötig, die aufgeworfenen Fragen vorab und explizit zu beantworten und »auszuhandeln«. Auch die Rahmung geschieht zusammen mit all dem, was im Gespräch noch geschieht, ohne dass wir davon viel Notiz nehmen. Versuche mit dekontextualisierten Ausschnitten aus Transkripten, die keine expliziten Hinweise darauf enthalten, aus was für einem Kontext die Ausschnitte stammen, zeigen, dass Versuchspersonen relativ schnell erschließen können, um was für einen Typ von Aktivität es sich handelt, wenn sie die entsprechenden, z.T. auch sehr kurzen Transkriptionsausschnitte vor sich haben. Unser Sprechen ist also auch mit Bezug auf diese Aufgabe eine starke Ressource: Es ist, wenn man so will, voll von Kontextualisierungshinweisen, die wir auf unterschiedlichsten Ebenen erzeugen und beim Zuhören wie selbstverständlich auswerten. Man sieht der Sprache deshalb in vielen Fällen noch den Kontext an, aus dem sie herausgerissen wurde. Sprechen-und-Zuhören lernen heißt insofern immer auch: Aktivitätstypen kennenlernen oder – in einer klassischen Terminologie – zu lernen, wie man Dinge mit Worten tut.¹⁶⁰ Es bedarf deshalb in der Regel gar keiner bewussten Reflexion, um das, was man mit Worten tut, entsprechend einzuahmen – weder als Sprecher:in noch als Hörer:in. Nur in Zweifelsfällen und wenn etwas schiefgeht, mögen wir tatsächlich unser Gegenüber fragen: »Was geht hier eigentlich gerade vor?« Dass das so selten geschieht, hat auch mit der Auswertung weiterer Ressourcen zu tun, die in unserer Darstellung schon angeklungen sind. Die für eine Gesellschaft besonders wichtigen Aktivitätstypen manifestieren sich nicht nur in der Sprache (z. B. auch als »Kommunikative Gattungen« und »Traditionen des Sprechens«)¹⁶¹ und in der Körperlichkeit (bestimmter Bewegungen und »Aufführungen«). Sie hinterlassen ihre Spuren auch an den Orten, an denen sie typischerweise stattfinden. In dem Maße, in dem sich solche Spuren verfestigen und zum mehr oder weniger gezielt aufgesuchten ortsfesten Bestandteil einer Praxis werden, entstehen allmählich eigene »Interaktionsarchitekturen«: materiale Sedimente der Lösung kommuni-

kativer Anforderungen.¹⁶² Von hier aus führt ein direkter Weg zu den gebauten Räumen und Funktionsgebäuden der modernen Gesellschaft, in und mit denen die vielen Formen institutionalisierter und organisierter Interaktion ein Zuhause gefunden haben. Von der Aufgabe der Rahmung her gesehen sind solche Interaktionsarchitekturen ausgesprochen leistungsfähige Ressourcen: Sie materialisieren feingliedrige Konfigurationen für gesellschaftlich ausdifferenzierte Kopräsenzkonstellationen und können als solche interaktiv in Anspruch genommen werden. Rahmung geschieht dann beispielsweise dadurch, dass sich die Beteiligten an einem bestimmten Ort (zu einer bestimmten Zeit) versammeln, z.B. im Hörsaal, und schon dadurch (vor-)klären, was zwischen ihnen demnächst ablaufen soll, z.B. eine Vorlesung. Natürlich kann die Interaktionsarchitektur nicht vorwegnehmen (determinieren), was für eine Art von Interaktion in ihr anlaufen wird, aber sie kann es hochgradig erwartbar und wahrscheinlich machen, dass ein bestimmter Aktivitätstyp von den Beteiligten als naheliegend erlebt und behandelt wird. Für das Zusammenfallen von Interaktion und Organisation, Anwesenheit und Mitgliedschaft,¹⁶³ ist die »richtige« Rahmung offenkundig von großer Bedeutung, damit die Beteiligten wissen, auf was sie sich gerade wechselseitig einlassen. Die Architektur trägt wesentlich dazu bei, diese Rahmung auf denkbar ökonomische und effektive Weise herzustellen. Intermittierend zwischen Interaktion und Gesellschaft, tritt die Organisation als spezifisch ausgeformte Interaktionsarchitektur in Erscheinung.

Wie die Positionierung taucht auch die Rahmung als Aufgabe in der klassischen gesprächsanalytischen Einführungsliteratur in der Regel nicht auf, obwohl sie in Goffmans Arbeiten sehr prominent vertreten ist.¹⁶⁴ Es ist in diesem Zusammenhang vielleicht auch kein Zufall, dass die Konversationsanalyse in ihrer Frühzeit um institutionalisierte Kommunikation einen Bogen gemacht hat, weil sie davon ausgegangen ist, dass die sie interessierenden Phänomene (z.B. des »turn taking«) dort nur in einer gegenüber dem Alltag verzerrten (z.B. vorstrukturierten) Form anzutreffen seien. Hier macht sich womöglich noch der Mythos Kopräsenz in der Annahme eines Ur- und Reinzustandes »natürlicher« Kommunikation bemerkbar. Diese Engführung ist in der einschlägigen empirischen Forschung inzwischen längst aufgegeben. Institutionalisierte Kommunikation ist zu einem eigenständigen Forschungsgegenstand geworden. Die Frage, wie sich dieser spezielle Kontext in der Interaktion bemerkbar macht, wie Kopräsenz durch Organisation und Institution konfiguriert werden kann, ist vielfältig dokumentiert und analysiert. Das Phänomen der Rahmung zeigt sich empirisch entsprechend vielschich-

tig: etwa als spezielle Art des Fragenstellens,¹⁶⁵ als Versuch, eine szenische Erzählung in einem Gespräch beim Sozialamt unterzubringen¹⁶⁶ oder darin, wie ich in einem Gespräch die Thematisierung von Sorgen und Nöten meines Gegenübers hervorlocke und darauf eingehe.¹⁶⁷ Diese Liste ließe sich mit Blick auf die gesprächsanalytische Untersuchung »institutioneller Kommunikation« nahezu endlos fortsetzen. Es dürfte aber auch so deutlich werden, dass und wie sich die Erledigung der Rahmung auf das Sprechen-und-Zuhören in Gesprächen auswirkt.

VII.

Wir hören mit dem Aufhören auf und kommen damit abschließend zu einer Aufgabe, die sich zwangsläufig ergibt, nachdem Kopräsenz einmal zustande gekommen und Interaktion einmal angelaufen ist. Die Anforderung der Beendigung ergibt sich allerdings nicht schon zwangsläufig aus der Zeitlichkeit und Episodenhaftigkeit der Interaktion mit Anfang und Ende, also daraus, dass Interaktion eben typischerweise irgendwann ausläuft. Es wäre ja nicht ausgeschlossen (und passiert vielleicht auch gelegentlich), dass die Beteiligten eben irgendwann »Abwesenheit wählen«, d.h. von einem Moment auf den anderen auseinandergehen und damit Kopräsenz auf einen Schlag auflösen. Auch wenn das aus praktischen Gründen nicht ganz einfach vorstellbar ist (man kann sich nicht einfach in Luft auflösen), zeigen die Erfahrungen mit auf Telekopräsenz beruhenden Videokonferenzen, worum es geht; hier kann es tatsächlich passieren (und passiert auch zuweilen), dass Teilnehmende warum auch immer auf einen Schlag nicht mehr dabei sind und der Bildschirm »einfriert« oder abschaltet. Auch diese vermeintliche Ausnahme bestätigt allerdings die Regel: Das plötzliche Wählen von Abwesenheit wird unter den nach wie vor Anwesenden typischerweise auf eine technische Unzulänglichkeit zurückgeführt (Bist du noch da?), als eine plötzliche Ohnmacht gegenüber höherer Gewalt. Diese Praxis verrät, dass das Aufhören tatsächlich eine interaktive Anforderung darstellt für die, die noch dabei sind, dass, anders gesagt, das Ende der Interaktion noch ein Teil der Interaktion ist. Für die Beendigung gilt so gesehen auch das, was wir bereits für die Eröffnung festgestellt haben: Es gibt eine eigenständige Sequenz für den Ein- und Austritt in die Interaktion hinein und aus der Interaktion heraus. Die Ränder der Interaktion werden interaktiv ausgestaltet mit und in Sequenzen, die ihrerseits eröffnet und beendet werden müssen. Das Problem des Anfangens und des Aufhörens wird sozusagen interaktiv verdoppelt und damit in die Interaktion

hereingeholt. Die Grenze zur Außenwelt dessen, was *noch* nicht oder nicht *mehr* Interaktion ist, wird dadurch interaktiv bearbeitet, dass sie zu einer Binnengrenze gemacht wird. Es ist deshalb zutreffend zu sagen, »dass die Gesprächsteilnehmer bei der Beendigung eines Gesprächs komplementäre interaktive Aufgaben bearbeiten müssen wie am Beginn«. ¹⁶⁸

Zu diesen Aufgaben gehört zunächst, »in spiegelbildlicher Reihenfolge«, ¹⁶⁹ die Einleitung der Abschlussphase, mit der sich die Beteiligten z. B. aufzeigen, dass die thematischen Erwartungen, die mit der Episode verknüpft und aufgebaut wurden, auch eingelöst und zudem keine dringenden oder aus welchen Gründen auch immer anliegenden Themen unbearbeitet geblieben sind. Die Einleitung der Abschlussphase ist deshalb prinzipiell offen für Nachträge und Ergänzungen, generell für bislang noch nicht Thematisiertes bzw. Erwähntes, das aber sehr wohl erwartbar thematisierbar und erwähnbar wäre (»unmentioned mentionables«). ¹⁷⁰ Komplementär zur Eröffnungsphase gehört weiter auch zur Beendigungsphase die ritualisierte Würdigung der Episode in Form der Bekundung der Achtung der Gesprächspartnerin oder des Gesprächspartners. Darin eingebettet sind Paarsequenzen der Verabschiedung, die bis an den Rand der (verbalen) Interaktion reichen und diese tatsächlich abschließen (»terminal exchange«). ¹⁷¹ Zuvor sind natürlich noch weitere Austauschpaare möglich wie Wunsch-Dank-Sequenzen, rückblickende Versicherungen des Wertes der Episode oder auch vorausschauende Ausblicke auf späteres Wiedersehen. Man weiß, wie sehr sich solche Achtungsrituale am Ende von Gesprächen in die Länge ziehen können – und über Umwege dann (doch noch) auf weitere Themenpotentiale stoßen können, die die Abschlussphase bis auf weiteres wieder außer Kraft setzen. Es gehört zu den Eigentümlichkeiten der Gesprächsdynamik, dass die Gesprächsbeendigung zwar mehr und mehr erwartbar und wahrscheinlich gemacht, aber tatsächlich nicht vorweggenommen werden kann. Dies gilt, solange wahrgenommen werden kann, dass wahrgenommen wird. In der Regel markiert auch der letzte Austausch deshalb noch nicht das Ende der Interaktion. Dazu bedarf es vielmehr der Auflösung wechselseitiger Wahrnehmung (»disengagement«), ¹⁷² die typischerweise erst nach dem Sprechen-und-Zuhören mit der Abwendung der Körper und dem Auseinandergehen einsetzt. Dem *späten* Einsetzen der Verbalität im Rahmen der Eröffnung entspricht so gesehen in vielen Fällen ein *frühes* Verebben der Verbalität im Rahmen der Beendigung. Nur unter telekopräsenten Bedingungen scheint es jedenfalls möglich, sich in einem Zuge mit dem »terminal exchange« aus der Interaktion auszuklinken. Man denke auch daran, dass die körperliche Konfiguration der Beteiligten oftmals

nicht einfach aufgelöst, sondern zunächst zugunsten körperlicher Verabschiedungszeremonien (z.B. Händedruck oder Abschiedsumarmung) verändert wird. Das kann überlappend zur Verabschiedung, aber auch sequentiell, d.h. nach dem Abschiedsgruß, erfolgen.

Die Unwahrscheinlichkeit der Interaktion ergibt sich im Fall der Beendigung unmittelbar aus der erfolgreichen Erledigung der anderen Aufgaben, insbesondere aus der Erledigung des Sichabwechslens: *Nach* dem Sprecherwechsel, kann man mit Fug und Recht sagen, ist *vor* dem Sprecherwechsel! Mit dieser Entdeckung im Hintergrund hatte die frühe Konversationsanalyse paradigmatisch am Beispiel der Gesprächsbeendigung die Unwahrscheinlichkeit der Interaktion entdeckt und analytisch ausgebeutet. Wie sollen die Beteiligten, wenn ein Gespräch erst einmal in Gang gekommen ist, gemeinsam und gleichzeitig zu einem Ende kommen, wenn dieses nicht von außen z.B. durch einen Gong signalisiert wird? Und selbst wenn es einen Gong oder ein anderes Signal von außen gibt, bricht die Interaktion daraufhin ja nicht gleich ab, sondern muss – gemäß dem, was wir eingangs ausgeführt haben – geregelt zu einem eigenen Ende geführt werden. Bei der Synchronisation kann ein Gong behilflich sein, aber er löst das Problem noch nicht. Und erst recht, wenn wir nicht bereits an institutionalisierte Veranstaltungen mit vorher feststehendem zeitlichen Limit denken, ergibt sich aus der Art und Weise, wie das Sichabwechseln rekursiv gelöst wird, ein massives Folgeproblem: Wie sollen es die Beteiligten schaffen, gemeinsam an einen Punkt in der Interaktion zu gelangen, an dem erstens keine neuen Themen mehr eingeführt und alte Themen nicht mehr weiterentwickelt werden und zweitens der Mechanismus des Sprecherwechsels außer Kraft gesetzt ist, ohne dass dies als einseitiges Schweigen oder plötzlicher Abbruch der Interaktion verstanden wird?¹⁷³ Man kann an dieser Stelle anschaulich nachvollziehen, wie der Einblick in die Probleme des Sprecherwechsels wie von selbst ein weiteres Problem entstehen lässt. Die Rekonstruktion der Art und Weise, wie das Sichabwechseln praktisch gelöst wird, läuft im Grunde darauf hinaus, einen Sprecherwechsel immer wieder wahrscheinlich zu machen und damit Schweigen der Tendenz nach zum Verschwinden zu bringen (auch wenn es natürlich situations- und kontextabhängig immer wieder auch längere Gesprächspausen geben kann).¹⁷⁴ Es muss deshalb auch eine Vorkehrung dafür geben, den Sprecherwechsel nicht nur zeitweilig (wie z.B. beim Erzählen- und -Zuhören),¹⁷⁵ sondern dauerhaft zum Ende der Episode hin außer Kraft zu setzen. Die analytische Fruchtbarkeit (wohlmeinend könnte man auch sagen: der empirische Beweis) der konversationsanalytischen Vorgehensweise

zeigt sich also darin, dass sie mehr und mehr Unwahrscheinlichkeiten der Interaktion zutage fördert und damit mehr und mehr an Verfahren zur Lösung dieser Unwahrscheinlichkeiten ans Licht bringt. Der letzte Austausch (»terminal exchange«) ist deshalb nicht nur ein Achtungsritual, und er ist auch nicht einfach nur zeitlich und im Nachhinein gesehen der *letzte* Austausch (gewesen). Der »letzte Austausch« ist vor allem ein Mechanismus, um aus dem Sprecherwechsel ohne Schaden wieder herauszukommen. Er ist insofern eine echte konversationelle Errungenschaft und Institution: Mit dem erfolgreichen Vollzug des letzten Austausches wird das anschließende Schweigen und Auseinandergehen erwartbar gemacht. Man weiß diesen Mechanismus nur zu schätzen, wenn man zuvor auch die gesamte Problematik des Sprecherwechsels vor Augen hat. Im Übrigen muss nicht nur das Sichabwechseln ein geregeltes Ende finden. Auch das Beitragen und die damit verbundene Themenprogression hören nicht von selbst einfach auf. Die Einleitung der Abschlussphase geht deshalb häufig einher mit Themaabschlusshinweisen (und womöglich mit Themawiedereinführungshinweisen), worauf wir bei der Besprechung der Themenorganisation bereits hingewiesen hatten.

Mit dem Abschluss der Besprechung des Beendens beenden wir auch unseren kursorischen Durchgang durch die zentralen Problem- und Aufgabenstellungen verbaler Interaktion. Alle diese Probleme lassen hervortreten, worin die genuine Unwahrscheinlichkeit der Interaktion besteht. Der Blick auf die empirische Bearbeitung der mit diesen Problemen verbundenen Aufgaben zeigt, wie dabei die Ressourcen der Interaktion – allen voran die natürliche Sprache – genutzt werden, um Interaktion im Einzelfall immer wieder möglich und wahrscheinlich werden zu lassen. In ihrer Konfiguration als verbale Interaktion hat sich Kopräsenz zu einem komplexen »multitasking system« entwickelt, mit dem durch das Anfangen und Beenden Episoden des Gesellschaftsvollzugs gebildet werden, innerhalb derer sich die Beteiligten als Sprechende und Zuhörende abwechseln und mit ihren Beiträgen ausgewählte Themen bearbeiten. Über die Realisierung von Anwesenheit hinaus zeigen und positionieren sie sich in den für ihr Beisammensein gerade relevanten Aspekten ihrer sozialen Identität und rahmen das laufende Geschehen als (Mit-)Vollzug von Gesellschaft. Interaktionsfähigkeit erweist sich dabei als ein komplexes Kürzel für die Bedingungen der Möglichkeit, an dieser Form von Kopräsenz teilnehmen zu können. Dabei haben wir am Beispiel der in vielerlei Hinsicht basalen Aufgabe der *Situierung* bereits Aspekte des evolutionären Wandels von Kopräsenz angedeutet: Bevor es Sprache gab, gab es bereits Phänomene von Ko-Orientierung, Ko-Ordination und Ko-Operation. Im folgenden Ka-

pitel wollen wir diesen Fokus auf Kopräsenz im Wandel aufnehmen, die Blickrichtung umdrehen und abschließend ausgewählte Zukunftsperspektiven thematisieren.

