

5 Kopräsens im Wandel – zum Abschluss ein Blick nach vorn

Wenn in diesem Essay von Kopräsens die Rede ist, geschieht das auch unter dem Anspruch, zumindest punktuell über eine rein synchrone Perspektive hinauszugehen und nach Kopräsens im Wandel zu fragen. Es hat das nicht primär mit einem historischen oder evolutionstheoretischen, geschweige denn mit einem prognostisch-prophetischen Interesse zu tun. Vielmehr geht es darum, den Gegenstand so zu bestimmen und zu fassen zu bekommen, dass er nicht in seinen rezenten Erscheinungsformen aufgeht und wir dem auf den Leim gehen, was in dieser Schrift mit dem Mythos Kopräsens verbunden wird.¹ Auch wenn Kopräsens das soziale Zuhause von Sprache ist und durch Sprache selbst maßgeblich geprägt wird, braucht es deshalb einen hypothetischen Rück- und Seitenblick auf Kopräsens *ohne* Sprache. Nur so kann in Ansätzen plausibel werden, wie und warum sich Sprache in diesen Interaktionszusammenhang so erfolgreich und so folgenreich einnisten konnte. Es gilt diese Perspektive nicht nur mit Blick auf die Genese und bisherige Entwicklung von Kopräsens, sondern auch mit Blick auf die weitere Entwicklung und Zukunft von Kopräsens. Es zeichnet sich an dieser Stelle zwar (noch) kein Jenseits verbaler Interaktion ab. Aber es gibt andere Entwicklungen, die es lohnen, abschließend näher betrachtet zu werden. Zwei dieser Entwicklungen wollen wir in diesem letzten Kapitel herausgreifen. Beide tragen auf ihre Weise zur Dekonstruktion des Kopräsensmythos bei. Da ist zunächst die Frage danach, was wir auf der Seite der Akteure und Akteurinnen voraussetzen müssen, damit sie als kopräsent Beteiligte für Interaktion überhaupt in Betracht kommen können. Es mehren sich die Anzeichen, dass es sich dabei nicht nur um Menschen im Vollsinne entwickelter und uneingeschränkter Sensomotorik und Kognition handeln muss, wie es der »Agency«-Idealisierung vom »isolierten vollkompetenten Akteur«² entspricht. Unter bestimmten Bedingungen können bekanntlich auch Kleinkinder, Menschen mit fortgeschrittener Demenz, Haustiere, Gott-

heiten und Geister sowie Gespenster »kopräsent« mit Menschen sein, mehr und mehr in jüngster Zeit auch Roboter und künstlich intelligente Technik. In der einschlägigen Forschung spricht man diesbezüglich (noch) von Grenzfällen und -bereichen des Sozialen. Aber was sind das für »Bedingungen«, die wir an dieser Stelle ansetzen müssen, und wie zeigen sich darin Anzeichen für einen Wandel von Kopräsenz? Darauf wollen wir gleich einen Blick werfen, ohne dass wir dabei systematisch und mit großer Tiefenschärfe auf einen sich an dieser Stelle rasant entwickelnden Forschungsstand eingehen können.

Die andere Entwicklung, die wir abschließend in Augenschein nehmen wollen, greift ein Phänomen auf, mit dem wir bereits viel vertrauter sind, obwohl es erst mit der Pandemie in den letzten Jahren vermehrt ins Bewusstsein gedrungen ist. Es ist das Phänomen, dass wir offenbar auch dann kopräsent sein können, wenn wir nicht nah beieinander, sondern einander räumlich fern sind. Das Telefon ist der inzwischen schon recht altmodisch anmutende Prototyp dieser Art von Telekopräsenz. Aber was genau ist damit gemeint, wenn wir hier in Anlehnung an einen Vorschlag aus der Sekundärliteratur von Telekopräsenz sprechen? Und gilt das, was für das Telefonieren gilt, abgesehen von der Verfügbarkeit einer Videoaufzeichnung, in gleicher Weise für die Beteiligung an Videogesprächen und -konferenzen, wie wir sie in den letzten Jahren in großer Zahl praktiziert haben? Das ist nicht nur ein viel diskutiertes Zeitgeistthema in den Feuilletons. Es ist für uns relevant, weil telekopräsente Interaktion das Raum-Zeit-Apriori des Kopräsenzmythos in Frage stellt, nach einem *sozialen* anstelle eines *physischen* Anwesenheitsbegriffs verlangt und Interaktion als eine gesellschaftlich robuste und technisch-medial wandlungsfähige Sozialform hervortreten lässt. Mit einem Blick auf solche Anzeichen für zukünftige Entwicklungen von Kopräsenz, die mit Refigurationen durch Kommunikationstechnologie zu tun haben, wollen wir den vorliegenden Essay beschließen.

Interaktionsfähigkeit: Wer kann kopräsent sein?

Vielfach sind die Konversationsanalyse und die Gesprächsforschung wie selbstverständlich davon ausgegangen, dass es sich bei den Interaktionsbeteiligten um Menschen handelt, oftmals und typischerweise um erwachsene, nicht speziell in ihren Wahrnehmungs-, Bewegungs- und Handlungsmöglichkeiten eingeschränkte Menschen. Das heißt nicht, dass es nicht auch Untersuchungen von Interaktionen mit Kleinkindern oder »Menschen mit

Beeinträchtigungen« geben würde.³ Gemeint ist, dass über die Frage, wer überhaupt als Interaktionspartner:in in Frage kommt und von welchen Voraussetzungen man bezüglich der Teilnahme an Interaktion auszugehen hat, vielfach nicht reflektiert worden ist. Interaktionsfähigkeit ist bis heute kein etabliertes Konzept in der Konversationsanalyse und Gesprächslinguistik. Wahrscheinlich kommt es der Forschungspraxis recht nahe, von der Annahme auszugehen, dass es sich bei den beteiligten Akteurinnen und Akteuren um Menschen handelt, die als solche fähig sein müssen, sinnhaft-kommunikativ zu handeln. Mit Alfred Schütz und mit Blick auf die Spezifik sinnhaft-kommunikativen Handelns kann man das z.B. so verstehen, dass die »grundsätzlichen Voraussetzungen für kommunikatives Handeln« in einer »Idealisierung der Reziprozität der Relevanzsysteme und [...] der Austauschbarkeit der Standpunkte« bestehen, vor deren Hintergrund eigenes Handeln entworfen, geprüft und umgesetzt wird.⁴ Mit Michael Tomasello und mit Blick auf den Menschen und die Menschwerdung kann man das so verstehen, dass »kooperative Akteure« durch gattungsspezifische Kompetenzen ausgezeichnet sind:

Es bedarf (1) der kognitiven Fertigkeiten zur Erzeugung gemeinsamer Intentionen und Aufmerksamkeit (und anderer Formen eines gemeinsamen begrifflichen Hintergrunds) mit anderen; und (2) sozialer Motivationen, anderen zu helfen und Dinge mit anderen zu teilen (und gegenseitige Erwartungen hinsichtlich dieser kooperativen Motive zu bilden).⁵

Die Liste solcher Kompetenzen ließe sich fortsetzen.⁶ Aber braucht es überhaupt eine auf diese Weise handlungstheoretisch und/oder anthropologisch aufgeladene Vorabbestimmung der Teilnahmebedingungen an Interaktion? Oder könnte man gerade auch mit der ethnomethodologischen Konversationsanalyse davon ausgehen, dass sich die Interaktion schon die Art von Interaktionsfähigkeit (be)schafft, die es braucht? Und was genau ist mit Interaktionsfähigkeit dabei gemeint? Und wie könnte sie in und mit Interaktion hergestellt werden? Mit Fragen wie diesen sind wir mitten im Thema dieses Kapitels.

Wir werden dies Thema im Folgenden in drei Schritten erarbeiten. Zunächst wollen wir uns mit der verbreiteten Annahme beschäftigen, dass es auf Akteursseite menschliche Handlungsfähigkeit als Teilnahmebedingung für Interaktion benötigt, und erläutern, warum wir mit einer solchen Annahme nicht weiterkommen. Im Kern geht es darum, dass man der Ei-

gengesetzlichkeit der Interaktion nicht gerecht wird, wenn man ihr die Frage der Teilnehmerschaft gleichsam aus der Hand nimmt. Es geht, um eine viel zitierte Formulierung von Goffman aufzunehmen, auf die wir noch zurückkommen, nicht um »Menschen und ihre Situationen«, sondern um »Situationen und ihre Menschen«. ⁷ Wir geben damit, abstrakter gesprochen, den Grundbegriff der Handlung auf und ersetzen ihn durch den Grundbegriff der Kommunikation. ⁸ Löst sich die titelgebende Frage, wer (überhaupt) kopräsent sein könne, also in Luft auf, weil Interaktion, salopp formuliert, frei darin ist, selbst zu definieren, was es auf der Seite der Beteiligten braucht? Offenkundig ist es nicht so einfach. Im zweiten Schritt wollen wir deshalb zeigen, dass und wie Interaktion mit Interaktionsfähigkeit selbst genau die Teilnahmebedingungen kreiert, die es braucht und die ihren Ausgangspunkt in Anzeichen für Wahrnehmungsfähigkeit und Adressierbarkeit haben, mit denen mindestens zwei Interaktionsbeteiligte in Erscheinung treten können. Diese Anzeichen können unter geeigneten Bedingungen in komplexen Erwartungen an »Personen« und »Akteurinnen« bzw. »Akteure« als »Adressen« der Kommunikation kondensieren und schließlich in weitreichenden anthropozentrischen Normalitätsvorstellungen idealisierter »agency« zum Ausdruck kommen (»theory of mind«). Im Routinefall der Etablierung von Kopräsenz bleibt die Konstruktion von Interaktionsfähigkeit unter der Oberfläche als implizite Hervorbringung. Mehr Evidenz liefern deshalb Situationen, in denen die Normalitätsvorstellungen aus irgendwelchen Gründen brüchig werden (im Umgang z.B. mit Menschen mit Erkrankungen, die sich auf Wahrnehmungsfähigkeit auswirken) oder in denen Interaktion ihre Grenzen erweitert (z.B. in Praktiken der Als-ob-Behandlung von Abwesenden als Anwesenden oder von Kleinkindern und Haustieren als intentional handelnden Akteurinnen und Akteuren). Im dritten und letzten Schritt wollen wir einen Blick auf solche Situationen werfen, weil sie einerseits zeigen, wie robust Interaktion sein kann, und andererseits, welche Art(en) von Grenzfällen der Transzendierung der Grenzen von Kopräsenz unsere moderne Gesellschaft womöglich (noch) zulässt.

I.

Wenn gegenwärtig viel über »Grenzbereiche der Kommunikation« und »Grenzen des Sozialen« ⁹ diskutiert wird, steht dahinter in der Regel die Annahme, dass Kommunikation *unter Menschen* den Normalfall und kommunikatives Handeln zwischen Menschen den Bezugspunkt bildet. Grenzbereiche erge-

ben sich dann in dem Maße, in dem von diesem Normalfall in irgendeiner Weise abgewichen wird und das vorausgesetzte Menschsein nicht oder nur eingeschränkt gegeben zu sein scheint. Dabei wird in der Regel nicht systematisch zwischen Interaktion, die auf Kopräsens beruht, und anderen Realisierungsformen von Kommunikation unterschieden, so dass auch die von uns eingeführten Fälle der Kommunikation unter der Bedingung von Lesbarkeit oder Benutzbarkeit, also Kommunikation mit und durch Schrift oder mit und durch Architektur, als Grenzbereiche anzusehen wären. Weder beim Lesen eines Textes noch beim Betreten eines modernen Funktionsgebäudes braucht es zwei Menschen, damit Kommunikation zustande kommt. Das einsame Lesen eines Textes und das einsame Betreten eines Funktionsgebäudes laufen ohne Kopräsens ab, stiften also keine Interaktion, wohl aber Kommunikation. Man muss dazu nicht Texte und Architektur als Quasiakteurinnen verstehen (und mystifizieren). Es reicht anzuerkennen, dass in diesen Fällen geeignete Erscheinungsformen auf den Plan treten, die (anstelle einer anderen Person) die Unterstellung nahelegen und hochgradig erwartbar machen, dass eine wie auch immer versteckte Mitteilungsabsicht vorliegt und eine Beobachterin oder ein Beobachter in dem Sinne versteht, dass sie oder er einen Unterschied zwischen Information und Mitteilung macht.¹⁰ In diesem Sinne ermöglicht die moderne Gesellschaft Kommunikation ohne Interaktion. Das gilt recht sicher auch dann, wenn Schrift in einer Gesellschaft nicht zur Verfügung steht. In dem Maße, in dem kopräsensbasierte Interaktionsepisoden materiale Spuren hinterlassen, die als solche (wieder-)erkannt werden können, dürfte es nahegelegen haben, solche Spuren nicht nur mehr oder weniger zwangsläufig und unwillkürlich zu hinterlassen, sondern auch mehr oder weniger bewusst und zielgerichtet zu produzieren (z.B. Malereien auf Felswänden). Zumindest mag man damit experimentiert haben, eine spätere Rezeptionssituation im Raum vorwegzunehmen und Formen zu hinterlassen, die als gelegte Spuren »verstanden« werden konnten – in dem wenig anspruchsvollen Sinn der Unterscheidung von Information und Mitteilung. Um Interaktion handelt es sich bei solchen frühen Rezeptionssituationen nicht, eher um erste sich abzeichnende kommunikative Alternativen zur Interaktion, die dazu angeregt haben mochten, auf diesem Wege weiter zu gehen. Sollte das je ein Grenzfall von Kommunikation gewesen sein, ist er es spätestens mit der Verbreitung von Schrift längst nicht mehr. Ungeachtet dieser Relativierung gilt »bewusstes und intentionales Handeln in Kopräsens« nach wie vor als gegebener Normalfall menschlichen Handelns.¹¹ In der Frage danach, wer überhaupt kopräsent sein kann und als interaktionsfähig

gelten kann, sind Vorstellungen menschlichen Handelns deshalb Ausgangs- und Fluchtpunkt zugleich.¹² Vorausgesetzt wird also so etwas wie *menschliche Handlungsfähigkeit* («Akteurskompetenz», «agency») als etwas, das Akteurinnen und Akteure mitbringen müssen, damit sie als Interaktionsbeteiligte in Frage kommen. Interaktion wird entsprechend häufig von vornherein als »zwischenmenschliches Handeln« verstanden, so dass sich die Frage, wer kopräsent sein kann, aus dieser Perspektive gar nicht stellt (abgesehen von Fällen, in denen das Vorliegen uneingeschränkter Handlungsfähigkeit offenkundig in Frage steht). Man stößt an dieser Stelle also typischerweise auf eine doppelte Engführung: auf eine »anthropologisch fundierte Humanzentrierung«¹³ mit »human-only restrictions«¹⁴ einerseits, der zufolge nur Menschen als Interaktionsbeteiligte in Frage kommen (Anthropozentrismus),¹⁵ und auf die Orientierung am Primat sinnhaften sozialen Handelns andererseits, der zufolge Interaktion aus Handlungen hervorgeht und besteht. Die Frage ist dann (nur noch), wer (unter welchen Bedingungen) sinnhaft sozial handeln kann und was im Einzelnen derartiges Handeln ausmacht und was man dabei voraussetzen muss. In dieser Prämisse, die die Grundlage der »Agency«-Idealisierung bildet,¹⁶ kommt aus unserer Sicht eine »substantialistische« Verengung zum Ausdruck; die Frage nach Interaktionsfähigkeit wird mit substantiellen Merkmalen menschlicher Handlungsfähigkeit beantwortet, die der Interaktion vorausgehend bei denen vorausgesetzt werden müssen, die sich auf eine Interaktion einlassen. Wir wollen nicht bestreiten, dass es in vielen Fällen, mit denen wir uns beschäftigen, tatsächlich handlungsfähige Menschen sind, die als Interaktionsbeteiligte fungieren. Wir wollen aber bestreiten, dass man auf diese Weise von außerhalb der Interaktion auf die uns interessierenden Teilnahmebedingungen für Interaktion zugreifen muss und sollte. Natürlich haben Goffman und andere Kopräsenztheoretiker der ersten Stunde bei ihren Überlegungen an Menschen (Personen wie dich und mich) als Beteiligte gedacht – vielleicht noch an Menschen am Telefon, aber eben doch an Menschen, wenn auch nicht immer nur an »normale« Menschen.¹⁷ Goffman spricht immer wieder vom »Zusammenkommen von Personen«, um auf die »zeitlich begrenzten Interaktionen« zu kommen, »die daraus hervorgehen« und die ihn ganz vorrangig interessieren. Worauf er hinauswill, sind eben nicht die Personen oder »Individu[en]«, sondern die »Beziehungen zwischen den Handlungen«:

Ich setze voraus, dass der eigentliche Gegenstand der Interaktion nicht das Individuum und seine Psychologie ist, sondern eher die syntaktischen Bezie-

hungen zwischen den Handlungen verschiedener gleichzeitig anwesender Personen. Da es aber die Handelnden sind, die das Grundmaterial liefern, wird es immer sinnvoll sein, zu fragen, welche allgemeinen Eigenschaften sie haben müssen, wenn diese Art von Beitrag von ihnen erwartet wird. Welches Grundmodell des Handelnden ist erforderlich, wenn wir ihn aufziehen, seinen Mitmenschen aussetzen sollen und das Ergebnis eine geordnete Interaktion sein soll? Welches Grundmodell benötigt der Forscher, um die Strategien antizipieren zu können, mit denen ein Individuum *qua* Interagierender erfolgreich sein oder zusammenbrechen kann?¹⁸

Wie Goffman wenig später betont, braucht es die Psychologie des Individuums nur »in ganz vereinfachter und verkürzter Form, um der soziologischen Analyse von Gesprächen, Jagdveranstaltungen, Banketten, Prozessen und Stadtbummeln zu genügen.« Vor diesem Hintergrund kann er seine Einleitung dann mit einem programmatischen Statement abschließen, in dem die Verlagerung des eigenen Interesses fulminant auf den Punkt gebracht wird: »Es geht hier also nicht um Menschen und ihre Situationen, sondern eher um Situationen und ihre Menschen.«¹⁹ Das, was das Menschsein ausmacht als Interaktionsteilnahmebedingung, ist deshalb von der Interaktion her (den jeweiligen »Situationen«, dem Individuum als »Interagierender« bzw. »Interagierendem« und den »Beziehungen zwischen den Handlungen«) zu gewinnen. Situationen, möchte man mit Goffman fortsetzen, schaffen sich die Menschen, die sie als Interagierende brauchen. Und je nach Situation braucht es dafür womöglich nicht einmal immer Menschen.

Ein erhellender Fall in diesem Zusammenhang ist der Anruf eines Voicebots von Google bei einem Restaurant oder Friseursalon, um einen Tisch bzw. einen Termin für einen Haarschnitt zu reservieren (»Hi, may I help you?« – »Hi, ehm, I'd like to reserve a table for Wednesday, the 7th«).²⁰ Wenn man der bis heute viel verbreiteten Geschichte glauben darf, haben die angerufenen Personen (im Friseursalon und Restaurant) nicht gemerkt, dass sie nicht mit einem Menschen aus Fleisch und Blut, sondern mit einem Roboter gesprochen haben – und das, obwohl die Reservation nicht einfach so klappt, sondern eine Komplikation auftritt, die auf beiden Seiten Nachfragen und Anpassungen erfordert. Das in etwa ist die medial vor einem Publikum in Szene gesetzte Pointe der Vorführung. Das Erhellende an diesem Fall ist nicht so sehr die womöglich gelungene Täuschung (und die Fragwürdigkeit der Authentizität der Beispiele). Was wäre z.B., wenn auch der oder die Angerufene auf der anderen Seite »nur« ein Voicebot gewesen wäre? Erhellend ist, wie wir mit

einem solchen Fall umgehen. Aus einer substantialistischen Perspektive wäre man wohl gezwungen, von einer Scheininteraktion zu sprechen, weil eine Seite der am Gespräch Beteiligten offenkundig nicht die Teilnahmebedingungen für Interaktion und Kommunikation im Sinne menschlicher Handlungsfähigkeit erfüllt. Anders sieht die Sache aus, wenn man sich nicht an das Vorwissen, sondern nur an den Verlauf des Anrufes hält. Das Nichtmenschliche oder auch nur irgendwie Problematische des Anrufenden wird in der Interaktion selbst nicht in einer Weise relevant, die zu nennenswerten Irritationen auf der Oberfläche führt. Die Stimme des Anrufenden klingt nach einer menschlichen Stimme am Telefon, der Anrufende spricht mit Anzeichen von spontan-ungeplanter Stegreifmündlichkeit (»ehm«), er grüßt an der richtigen Stelle zurück (»Hi«), formuliert sein Anliegen passgenau und antwortet thematisch (mehr oder weniger) kohärent auf sich erst im Verlauf ergebende Nachfragen. Mit all dem verhält er sich so, dass ein Anschein ausgelöst und aufrechterhalten werden kann, der der Angerufenen genügend Evidenz dafür bietet, dass sie wahrnehmen kann, dass sie wahrgenommen wird. Das wird in diesem Fall sicher durch die Situation am Telefon (Telekopräsenz) und die Vorhersagbarkeit des Ablaufs (technischer gesagt: des Sichabwechselns und des Beitragens zum Thema Reservation) begünstigt. Interessant(er) für uns aber ist nicht, dass »in Wirklichkeit« eine Maschine (nur) so tut, als sei sie eine Anruferin aus Fleisch und Blut, sondern dass auf der Ebene der Erscheinungsformen der Interaktion kein Anlass für entsprechende Irritationen geliefert wird, sondern dass erfolgreich so etwas wie Interaktionsfähigkeit demonstriert wird. Man könnte auf die Idee kommen, dass es in »natürlich-authentischer« Interaktion prinzipiell nicht anders ist: Die Interaktion hat – zumal am Telefon – nur das zur Verfügung, was sich die Teilnehmer:innen sinnlich wahrnehmbar präsentieren. Was genau »dahinter« steckt, ist für die Interaktion im Sinne der viel zitierten Blackbox unerheblich. Man könnte auch sagen: In dem Maße, in dem sich die Bedienung einer Maschine nicht mehr von einem Austausch mit einem Menschen unterscheiden lässt, fallen Benutzbarkeit und Anwesenheit/Kopräsenz als Kommunikationsbedingungen zusammen. Das gilt übrigens auch in umgekehrter Richtung: Ein Mensch, der sich aufgrund von Routine und Professionalisierung in immer wieder gleichen Abläufen wie eine Antwortmaschine verhält, trägt womöglich dazu bei, dass die Interaktion mit einem Menschen sich nicht mehr von der Bedienung einer Maschine unterscheiden lässt. Davon sind wir offenkundig noch sehr weit entfernt, aber Beispiele wie solche Anrufe zeigen bei aller Einschränkung (und bei allen Zweifeln an ihrer Authentizität), wie man sich solche Verschmelzungen

vorstellen muss. Letztlich dürfte es dann tatsächlich zu einer Nebensache werden, ob am anderen Ende (hinter dem Schalter oder am Telefon) noch ein Mensch oder »nur noch« eine Maschine sitzt, die spricht und aussieht wie ein Mensch. Von einer solchen Relativierung humanspezifischer »agency« im Alltag der Kopräsenz sind wir erst recht noch weit entfernt – sonst könnten Beispiele wie die zitierten nicht so populär und spektakulär in Szene gesetzt werden. Aber sind der und die, die bei einem Restaurant anrufen, um einen Sitzplatz zu reservieren, mit ihrer Orientierung an der zu erledigenden Aufgabe und ihrem vorhersagbaren Ablauf tatsächlich so weit entfernt von der vorprogrammierten Anrufmaschine? Braucht es für einen solchen Anruf so viel mehr an »human agency«? Und ist nicht das genau der Sinn von Goffmans Einleitungsschluss, dass es um die »Situationen und ihre Menschen« geht (und nicht umgekehrt)? Natürlich kann auch ein Anruf für eine Sitzplatzreservierung – wie der Kauf eines Tickets beim Bahnschalter – sehr schnell mehr und anderes werden als nur die Abarbeitung eines Anliegens. Aber in vielen Fällen ist er genau das – mit und ohne Menschen auf der anderen Seite. Das Beispiel zeigt insofern auch, dass auch »unter Menschen« das Menschsein als Teilnahmebedingung nicht statisch und gegeben, sondern flexibel, dynamisch und emergent ist. Auch in diesem Sinn gilt, dass es unter Kopräsenten so viel Menschsein gibt, wie situativ gerade gebraucht wird.

Wir ziehen aus Beispielen wie den gerade zitierten die Konsequenz, die intuitiv-selbstverständliche Orientierung an Menschen aus Fleisch und Blut als Interaktionsbeteiligten aufzugeben und abstrakter davon auszugehen, dass so etwas wie Interaktionsfähigkeit keine externe Bedingung der Interaktion ist, sondern mit und durch Interaktion selbst ins Spiel kommt und dabei auf eine flexible Weise als Beiprodukt von Kopräsenz emergiert – und zwar in der Form und Ausgestaltung, die es gerade braucht. So selbstverständlich die Orientierung an Menschen aus Fleisch und Blut als Interaktionsbeteiligten ist, so wenig ist sie theoretisch gedeckt. Deshalb müssen wir Interaktionsfähigkeit erstens abstrakter und zweitens von der Interaktion her konzipieren. Man kann sich das am Beispiel der Intentionalität kommunikativen Handelns klar machen, die wie oben illustriert oftmals zur Teilnahmebedingung für Interaktion gemacht wird. Schon aus dem hier vertretenen Kommunikationsbegriff geht hervor, dass eine solche Vorannahme nicht gedeckt ist. Kommunikation kommt nach Luhmann, an dessen Kommunikationstheorie wir uns orientieren,²¹ nicht durch das Vorliegen einer kommunikativen Intention zustande (ich will die Andere oder den Anderen auf mich aufmerksam machen), sondern dadurch, dass eine solche Intention (im Sinne der Unterscheidung von Infor-

mation und Mitteilung) durch eine Andere oder einen Anderen unterstellt werden kann (das Winken mit der erhobenen Hand bei gleichzeitigem Blickkontakt und Gruß lässt mir keine andere Wahl). In diesem Sinn sind der Gruß der oder des Anderen am Telefon und die Anliegenformulierung starke Auslöser für das Unterstellen einer Mitteilungsintention. Entscheidend ist dabei nicht die oder der Andere in ihrer oder seiner interaktionsexternen Substanz und Identität als Akteur:in, sondern ob etwas passiert, dass auf eine Mitteilungsintention schließen lässt. Das zu explizieren, was es an dieser Stelle braucht, verlangt eine Abstraktion von gewohnten Vorkommen (Menschen aus Fleisch und Blut) zugunsten der dahinterliegenden Merkmale von Interaktionsfähigkeit (s. dazu u. II. in diesem Abschnitt) – und wie diese Merkmale zum Ausdruck kommen, d.h. wahrnehmbar werden. Die Last der Erklärung wird also auf die Frage verschoben, unter welchen Umständen und aufgrund welcher Erscheinungsformen eine solche Mitteilungsunterstellung ausgelöst (und ihrerseits erwartbar gemacht) werden kann. Das ist u.a. eine der Errungenschaften von Sprache: Sie ist als Form so unwahrscheinlich, dass die Unterstellung einer Mitteilungsabsicht gar nicht mehr verhindert werden kann. Unter denen, die sprechen und zuhören können, macht Sprache das Anlaufen von Interaktion also unausweichlich. Wenn Sprache *nicht mehr* verfügbar ist, muss etwas anderes einspringen. Und solange Sprache *noch nicht* verfügbar war, muss es Vorläuferinnen und Vorformen von Sprache gegeben haben (wie Gestik und Mimik), die die Unterscheidung von Information und Mitteilung ermöglicht und nahegelegt haben. Offenbar waren es erst Menschen, die (aus welchen Gründen auch immer) damit angefangen haben, auf diese Weise Kopräsenz zu einer genuin sozialen Erfahrung zu machen. Die Evolution der Kommunikation verläuft deshalb über die Evolution der Formen des Kommunizierens – und geht damit Hand in Hand mit einer Evolution der sonst noch beteiligten Systeme (etwa von Leben und Bewusstsein). Zweifellos verlangen die Produktion und Rezeption solcher Formen (z.B. des Grüßens durch Winken mit der Hand) spezielle Fertig- und Fähigkeiten, die man z.B. im Vergleich von Menschenaffen und (jungen) Menschen sehr genau bestimmen kann.²² Aber die in diesen Erscheinungsformen emergierende Interaktion wartet nicht darauf, dass diese Fertig- und Fähigkeiten vorhanden sind, sondern entwickelt sich ko-evolutionär in einem Zug *mit* diesen Fertig- und Fähigkeiten. Worauf es also ankommt, sind die Erscheinungsformen des Kommunizierens, genauer: Erscheinungsformen von Kopräsenz. In diesem Sinn ist das Menschsein keine Voraussetzung für Interaktion, sondern nicht mehr als ein Kürzel für Interaktionsfähigkeit und als solche in dem für Kopräsenz relevanten Sinn ein Effekt

der Interaktion. In den Erscheinungsformen kommt zum Ausdruck, was es an Interaktionsfähigkeit braucht und wie diese weiter ausgebaut und entwickelt werden kann – es geht, um Goffman zu paraphrasieren, nicht um Menschen und ihre Interaktionen, sondern um Interaktionen und ihre Menschen. Nur so kann man in der Phylo- und Ontogenese nachvollziehbar machen, wie womöglich durch und mit Interaktion (z.B. durch spezielle Muster der Erwachsenen-Kind-Interaktion) so etwas wie die Menschwerdung mit und durch Kopräsenz zustande gebracht worden ist und nach wie vor zustande gebracht wird – anstatt das Menschsein im Vollsinn der Phylo- und Ontogenese vorzugeben und damit schon vorauszusetzen, was es gerade zu beschreiben gilt. Zugleich können wir analytisch einholen, was genau am Menschsein die Interaktion braucht, damit sie anlaufen kann, und wie daraus womöglich allmählich etwas entsteht, dass dem Menschsein, wie wir es kennen, ziemlich nahekommt. Onto- und Phylogenese zeigen zugleich, dass und wie Interaktion auch damit umgehen kann, Fähig- und Fertigkeiten zu kompensieren, die womöglich in Grenzbereichen der Kommunikation (noch nicht oder nicht mehr) vorhanden sind. Und man kann zeigen, dass und wie Interaktion, an der mindestens ein Mensch beteiligt ist, im Prinzip auch auf Bereiche ausgedehnt werden kann, die über die Gattung Mensch hinausgehen. Die Gattungsbeschränkung ist also nicht unbedingt in Stein gemeißelt, sondern im Prinzip in und mit Interaktion überwindbar.

II.

Folgt aus dem, was wir hier als Erzeugung von Interaktionsfähigkeit durch Interaktion postuliert haben, dass Interaktion keine Grenzen kennt, weil sie prinzipiell alles wie Beteiligte behandeln und z.B. Angehörige Wachkomapatienten so behandeln können, *als ob* sie Begrüßungen wahrnehmen und auf Fragen reagieren würden; oder Fernsehzuschauer auf die Begrüßung durch eine Fernsehmoderatorin hin zurückgrüßen, also die Figur im Fernsehen wie eine Person im gleichen Raum behandeln (parasoziale Interaktion);²³ oder Hundebesitzerinnen ihre Lieblinge wie sinnhaft handelnde »companions« behandeln, indem sie von Verhalten auf Intentionen schließen und diese kommunizieren? Wer es sich so einfach machen wollte, würde offenkundig dem auf den Leim gehen und das für bare Münze nehmen, was die Beteiligten selbst suggerieren und was womöglich nicht mehr als »Projektionen der Akteure« sind.²⁴ »Gespräche« mit Haustieren bleiben Selbstgespräche, auch wenn die Haustiere dabei behandelt werden mögen wie voll ausgewachsene Dialogpartner. Ein

Autofahrer, der während des Autofahrens »mit seinem Auto spricht« und es auf den vorgesehenen Tankstopp einstimmt (und dabei das Armaturenbrett streichelt),²⁵ ist jemand, der so tut, als ob er mit einer Maschine interagieren könnte – und bleibt doch jemand, der eine Maschine bedient und währenddessen etwas sagt, was nur er hören und verstehen kann (und mag er sich noch so verstanden fühlen). Dass in diesem Fall keine Interaktion zwischen Mensch und Maschine vorliegt (und auch kein Grenzbereich von Interaktion), beruht allerdings nicht auf der interaktionsunabhängigen Evidenz, dass Autos offensichtlich (noch) nicht sprechen und hören können und als Interaktionspartner prinzipiell nicht in Frage kommen. Die Evidenz kommt vielmehr aus dem, was das Auto macht und *nicht* macht und dass es keinerlei Hinweise für manifestierte Interaktionsfähigkeit(en) auf Seiten des Autos gibt. Im Gegensatz dazu lässt sich der Anruf des Voicebots im Restaurant (s.o. I.) systematisch auf seinen Beitrag zur Interaktion und die darin implizierten und suggerierte(n) Interaktionsfähigkeit(en) rekonstruieren. Wenn man es so betrachtet, zeigt sich auch, dass der Fall des- oder derjenigen, der oder die mit seinem oder ihrem Hund spricht, bereits komplizierter ist. Als dem Menschen vertraute Lebewesen sind Haustiere viel leichter als Interaktionsbeteiligte adressierbar, zumal grundlegende Formen der Situierung unstrittig sind. Beispiele wie diese legen es nahe, über Interaktionsfähigkeit als Bedingung und Errungenschaft von Kopräsenz weiter nachzudenken.

Zunächst ist festzuhalten, dass Interaktionsfähigkeit als Kategorie an die Stelle einer substantialistischen (typischerweise humanzentrierten) Prämisse tritt. Sie bezeichnet Bedingung *und* Errungenschaft von Interaktion zugleich. Auf der Ebene der an Interaktion beteiligten Systeme (Leben und Kognition, biologische und psychische Systeme) handelt es sich um Fähig- und Fertigkeiten, die z.B. mit Wahrnehmungsfähigkeit und sozialer Kognition zu tun haben und die man als Bedingung ansetzen kann, damit Kopräsenz emergieren kann: Damit wahrgenommen werden kann, dass wahrgenommen wird, muss Wahrnehmung selbst überhaupt bis zu einem bestimmten Grad wahrnehmbar sein, also z.B. gesehen werden können, wohin jemand sieht – was evolutionär eine Errungenschaft der Biologie des menschlichen Auges ist, die z.B. mit der Beweglichkeit des Augapfels und der Lederhaut (Sklera) zu tun hat. Und damit sind wir erst am Anfang der komplexen biologischen und kognitiven Fähigkeiten, die in und mit Interaktion als Ressource ausgenutzt werden und die aus Phylo- und Ontogenese des Menschen hervorgegangen sind und bis heute hervorgehen. Auf der Ebene der Interaktion selbst handelt es sich bei diesen Errungenschaften um anlass- und situationsspezifisch emer-

gente Hervorbringungen, die an der Oberfläche der Erscheinungsformen der Interaktion manifest werden (müssen). Wenn ein »normaler wacher Erwachsener« unter Anwesenden einschläft, ist er oder sie allenfalls auf eine sehr beschränkte Weise, wenn man so will, gleichsam *latent* kopräsent und für die anderen Anwesenden kaum noch interaktionsfähig. Ein Mensch, der in einem Restaurant einen Tisch vorbestellt, muss für diesen Anlass nicht sein gesamtes biologisch, sozial und kognitiv einzigartiges Repertoire an Kompetenzen abrufen, um in dieser Situation sein Ziel zu erreichen. Interaktion kann sich also situations- und anlassbezogen ohne weiteres auf die Hervorbringung basaler Formen von Interaktionsfähigkeit beschränken, die womöglich auch von einer stimm- und dialogfähigen künstlichen Intelligenz erbracht werden können, die Unterhaltung (verbale Interaktion) simulieren kann (auch ohne dafür über Bewusstsein oder Intentionalität zu verfügen). Für die Interaktion spielt das *in dieser Situation* keine Rolle. Was wir damit sagen wollen, ist: Als komplexes Bündel evolutionärer Errungenschaften von Leben und Kognition ist Interaktionsfähigkeit allgemein beschreibbar, als interaktive Hervorbringung tritt sie dagegen immer situations- und anlassbezogen, und d.h. flexibel in Erscheinung, gerade so, wie es für die Zwecke der fraglichen Interaktion relevant ist. Interaktionsfähigkeit ist deshalb für die Interaktion nicht einfach gegeben oder nicht gegeben, sondern emergent. In einer flüchtigen Begegnung unter Passantinnen und Passanten auf einem Bürgersteig oder in einer Bahnhofshalle manifestiert sich Interaktionsfähigkeit u.a. darin, Bewegungsabläufe und Gehrichtungen (»trajectories«) zu antizipieren und Andere auf eine Weise zu grüßen oder nicht zu grüßen, die der Art der unterstellten sozialen Beziehung entspricht, jedenfalls auf andere Weise als in einem Fußballspiel oder in einer Seminardiskussion über Gottesbeweise. Interaktionsfähigkeit ist deshalb grundsätzlich situativ zu bestimmen anhand der Erscheinungsformen der Interaktion, in denen sie sich fallspezifisch verkörpert. Dabei würde sich vermutlich ein Spektrum an Interaktionsfähigkeiten abzeichnen, das von sehr kurzzeitig-flüchtigen Momenten von Kopräsenz im Sinne der Ko-Orientierung und Ko-Ordination bis zur Entfaltung ausdifferenzierter Leistungen anspruchsvoller Ko-Operation reicht, für das über die gesellschaftliche Sozialisierung hinaus spezielles Training und fachliche Schulung notwendig sind. Dabei kommt allem, was mit Wahrnehmungsfähigkeit zu tun hat, aufgrund der für Kopräsenz konstitutiven Rolle reflexiver Wahrnehmung (Wahrnehmungswahrnehmung) offensichtlich eine besondere Bedeutung zu. Wer oder was auch immer in Interaktion treten will, muss in der Lage sein, zu sehen, dass er, sie oder es gesehen wird und dieses Sehen-des-Gesehen-

Werdens als mitgeteiltes Dasein wechselseitig zuordnen und zuschreiben können. Natürlich kann das Sehen des Gesehenwerdens auch durch andere Sinneswahrnehmungen ersetzt werden. Und wie wir schon mehrfach betont haben, reicht in vielen Fällen auch die Erwartbarkeit und Unterstellbarkeit reflexiver Wahrnehmung. Aber das setzt dann schon sehr viel an erfolgreich erprobtem, vielfach geübtem und routinehaft eingespieltem Kopräsenzmanagement voraus.

Mit der Bedingung der Möglichkeit reflexiver Wahrnehmung ist eine Art von Spiegelung der eigenen Wahrnehmung in der Wahrnehmung eines oder einer Anderen verbunden. »It takes two to tango.« Wahrnehmungsleistungen müssen also wechselseitig zugeordnet und zugeschrieben werden können. Das ist ein Erfordernis der Interaktion, die als eigengesetzlich-genuin soziales Geschehen nur anlaufen kann, wenn und indem im Akt reflexiver Wahrnehmung mindestens zwei Wahrnehmungsinstanzen als Interaktionsbeteiligte in Erscheinung treten. Dabei ist es mit Wahrnehmungsleistungen allein nicht getan. Schon die Anforderung der Ko-Orientierung, die in vielerlei Hinsicht als basal gelten kann, meint ja nicht nur die Wahrnehmung der oder des Anderen bzw. dass eine Andere oder ein Anderer da (bzw. hier) ist, sondern dass dieses Da- und Hiersein auch »mitgeteilt« wird. So entsteht ein wechselseitig geteilter Wahrnehmungsraum, in dem die Saat der Kopräsenz aufgehen und aufblühen kann. Dafür braucht es nicht unbedingt eine fix und fertig ausgereifte Absicht zur Mitteilung, sondern Formen, die es nahelegen, dass reflexive Wahrnehmung nicht zufällig passiert, sondern Ausdruck und Manifest der »gemeinsamen Aufmerksamkeit«²⁶ der Beteiligten aufeinander geworden ist, was damit nicht länger ignoriert oder für zufällig gehalten werden kann. Dieses Dasein in einem geteilten Wahrnehmungsraum manifestiert in einem grundlegenden Sinn Interaktionsfähigkeit und schafft gleichzeitig die Bedingung dafür, dass die Beteiligten anfangen können, sich als Interaktionsbeteiligte zu entdecken, zu erleben und zu behandeln, kurz gesagt: Interaktionsbeteiligte (und damit vielleicht auch: Menschen) zu werden.

Das hier beschriebene Erfordernis der Zuordnung und Zuschreibung von Wahrnehmungsleistungen an Interaktionsbeteiligte, in denen sich Interaktionsfähigkeit manifestiert, ist in der Systemtheorie als Problem der Adressabilität von Kommunikation formuliert worden. Damit Kommunikation (also auch Interaktion) zustande kommen kann, muss sie mit dem konstitutiven Akt der Unterscheidung von Information und Mitteilung so etwas wie mitteilungsfähige Instanzen nahelegen, um nicht zu sagen: erfinden. Dafür ist die Kommunikation auf »Adressen« angewiesen, mithilfe derer

Wahrnehmungsleistungen im Akt der Kommunikation zugeordnet und zugeschrieben werden und Mitteilungsunterstellungen überhaupt »ankommen« können. Der Ausgangspunkt ist also nicht ein handlungstheoretisch vorausgesetztes, handlungs- und mitteilungsfähiges Agens, sondern ein Erfordernis der Kommunikation. Entsprechend variabel und sozial konstruiert ist das, was für Interaktion als Adresse in Frage kommen kann:

Irgend jemand, irgend etwas muß ein Verhalten vorgeführt haben, an das angeschlossen werden kann, weil es als Mitteilungsverhalten zu deuten gewesen ist, in diesem Sinne also als eine Tat, ein Akt, eine Handlung. [...] [E]s ist die Kommunikation, die im Management ihrer Selbstsimplifikation Zurechnungspunkte erzeugt und ausarbeitet, die dann als handelnde (mitteilende) Personen erscheinen, dies dann in jener Evidenz, die die Jahrtausende alte Täuschung ermöglichte, Kommunikation werde durch die Subjekte betrieben.²⁷

Adressen werden so gesehen im Nachhinein durch und mit Kommunikation erst in die Welt gebracht. Sofern es sich um Interaktion (Kommunikation unter Anwesenden) handelt, geht mit der Zuordnung von Mitteilungen mit dem Aufkommen von Adressen, der »kommunikativen Verfertigung von Akteuren«,²⁸ die Zuschreibung von Interaktionsfähigkeit einher. Es ist wichtig zu sehen, dass das Agens dieser Zuordnungs- und Zuschreibungsaktivitäten nicht die beteiligten Akteurinnen und Akteure sind, sondern die Interaktion selbst. Wir müssen diesen Zuordnungs- und Zuschreibungsprozess von Interaktionsfähigkeit also, wie schon betont, auf der Oberfläche der Erscheinungsformen der Interaktion nachweisen. Dazu gehört auch die Ausfüllung von dem, was überhaupt als adressierbar in der Welt gelten kann. Auch das ist nicht durch die Natur der »Weltvorkommnisse« selbst schon vorentschieden, sondern muss in und mit Interaktion aufgenommen und bestätigt werden. Peter Fuchs spricht in diesem Zusammenhang von »Eigenverhältnis« und »Selbstreferenzialität« als entscheidender Voraussetzung für die Zuschreibung und Ausarbeitung von Adressierbarkeit:

Die unabdingbare Voraussetzung dafür ist, dass adressable Weltvorkommnisse (zum Beispiel Menschen, Bäume, Computer) ein Eigenverhältnis unterhalten, oder anders: dass ihnen Selbstreferenz unterstellt werden kann. [...]

Die Kommunikation kommt nicht umhin, darüber zu befinden, ob etwas in ihrer Umgebung als selbstreferenzfähig gelten darf (also als ein Subjekt, das

ein Verhalten vorführt, an dem Mitteilung und Information unterschieden werden können) oder als ein Objekt, das, weil es keinen Selbstbezug unterhält, allenfalls thematisierbar ist.²⁹

Wie aus diesem Zitat hervorgeht, wird auch darüber, ob jemand oder etwas »als selbstreferenzfähig gelten darf«, in der Kommunikation (darin eingeschlossen: Interaktion) entschieden. Hier ist der Ort, an dem entschieden wird, ob »Weltvorkommnisse« (wie »Menschen, Bäume, Computer«) zu möglichen Interaktionspartnern gehören, mit denen man sprechen kann, oder aber zu dem, was alles nicht dazu gehört und *worüber* man allenfalls sprechen kann. In der Formenwelt natürlicher Sprachen ist dieser Unterschied in der Markierung der »Gesprächsrollen« und der »Handlungsrollen« grammatisiert.³⁰ Letztlich geht es dabei, wie bei Fuchs dann weiter ausgeführt wird, um Inklusion und Exklusion, wie auch immer man Interaktionsfähigkeit losgelöst von konkreten Erscheinungsformen konkreter Interaktionen auszubuchstabieren versucht.³¹ Weiterführend erscheint an dieser Stelle das Konstrukt der Person, das als Kürzel für ein Bündel abstrakter Merkmale fungiert, unter dem Konkretisierungen von Interaktionsfähigkeit in der Kommunikation kondensieren können. Dabei haben wir es nicht mit Menschen aus Fleisch und Blut zu tun, sondern mit »Erwartungscollagen«, die von Fall zu Fall unterschiedlich und d.h. flexibel aktualisiert und realisiert werden können, d.h. zu Erwartungserwartungen und sozialen Strukturen werden können. Insbesondere mit der Konfiguration von Kopräsenz durch Sprache und dem damit verbundenen Aufkommen und dem Ausbau weiter ausgreifender interaktiver Aufgaben des Anfangens, Sichabwechselns, des Beitragens zu einem Thema, der Positionierung, der Rahmung und des Aufhörens steigen die Anforderungen an das Person-Sein und die damit in der Interaktion manifestierten Interaktionsfähigkeiten erheblich an.³² Personen sind, könnte man in Anlehnung an Goffman zusammenfassen, Interaktionsbeteiligte, die aus solchen immer komplexer werdenden Situationen hervorgehen und von ihnen geformt und geprägt werden.

III.

Einen Einblick in die mit der Manifestation erwarteter Interaktionsfähigkeit verbundenen Implikationen vermitteln Kontexte und Situationen, in denen der Personenstatus eines der potentiellen Interaktionsbeteiligten aus unterschiedlichen Gründen von Anfang an problematisch scheint oder aber allmäh-

lich problematisch geworden ist. In der sozialwissenschaftlichen Forschung ist mit Bezug auf dieses sehr heterogene Feld von »Grenzen« und »Grenzbe-reichen« der sozialen Welt die Rede. Dazu gehören, um besonders prominente Fälle herauszugreifen, die Kommunikation mit Kleinkindern und Menschen mit Demenz oder auch die Kommunikation mit Haustieren, nicht menschlichen Primaten und künstlicher Intelligenz. Weil es hier ganz allgemein darum geht, wie sich in und mit Gesellschaft Konstrukte von Interaktionsfähigkeit und (nicht menschlicher) »agency« herausbilden,³³ ist das Feld mit den schon genannten und den im Anschluss aufgeführten Fallbeispielen keineswegs abschließend erfasst. Z. B. lassen sich auch Phänomene religiös motivierter Kommunikation (mit Gott) in den Zusammenhang nicht menschlicher »agency« stellen. Zwar hat Niklas Luhmann die Frage klar verneint, ob es Kommunikation mit Gott gibt – aber diese Frage im Titel des fraglichen Aufsatzes leicht dahin verschoben, ob »unsere Gesellschaft Kommunikation mit Gott zu[lässt]«. ³⁴ Die Frage zur Kommunikation mit Gott wird damit zu einer in der Gesellschaft selbst beantworteten Frage. Es könnte also (möglich) sein, dass in und mit Interaktion »Gott« als Adresse erfolgreich etabliert wird. Und es könnte sein, dass spezielle Orte, spezielle Funktionen und Ämter, spezielle Praktiken und Rituale eingeführt worden sind, um diese Unwahrscheinlichkeit der Interaktion mit Gott interaktiv überwindbar(er) zu machen – ohne dabei jedes Mal auf ein »Wunder« angewiesen sein zu müssen. Es würde zu weit führen, diese zumal nicht leicht überschaubaren Forschungsbereiche an dieser Stelle systematisch aufzuarbeiten. Im Folgenden greifen wir deshalb nur sporadisch ein paar ausgewählte Beobachtungen anhand einiger Fälle heraus, um zu zeigen, wie sich an der Oberfläche der Interaktion der Umgang mit problematischer Interaktionsfähigkeit manifestieren kann und welche Auswirkungen sich speziell aus Mensch-Roboter-Kontakten für die Zukunft der Interaktion ergeben könnten.

Wenn man zunächst bei Menschen bleibt, kann die Problematik des Personenstatus entwicklungsbedingt in Erscheinung treten. Säuglinge und Kleinkinder sind offenkundig erst noch dabei, ihren Status als »Personen« mit voller Interaktionsfähigkeit zu entwickeln und zu erwerben. Entsprechend aussagefähig sind Erwachsenen-(Klein-)Kind-Interaktionen, die in ihrer Entwicklungsrelevanz inzwischen sehr gut beschrieben worden sind (s.u.). Viel spricht dafür, dass Interaktion und Kopräsenz in Ko-Evolution mit der Menschwerdung entstanden sind. Spezifische menschliche Fähig- und Fertigkeiten sind, wie wir betont haben, eine grundlegende Ressource der Interaktion, und sie tragen dazu bei, Interaktion maßgeblich zu konfigurieren.

Sofern wir uns überhaupt noch in frühmenschliche Interaktion einzufühlen vermögen, ist klar, dass wir dabei nicht an den anspruchsvollen Fall verbaler Interaktion mit einem ausgebauten Set interaktiver Aufgaben denken dürfen, sondern eher an grundlegende Formen von Ko-Orientierung, Ko-Ordination und Ko-Operation. Die Erwachsenen-(Klein-)Kind-Interaktion ermöglicht es, dass wir auch heutzutage noch dabei zuschauen können, wie mit und durch Interaktion fehlende Fähig- und Fertigkeiten kompensiert werden können und wie dabei Interaktionsfähigkeit nicht nur an der Oberfläche der Erscheinungsformen manifest, sondern auch gefördert und eingeübt werden kann. Es fängt damit an, überhaupt einen Wahrnehmungsraum gemeinsamer Aufmerksamkeit zu etablieren, was in der Kommunikation mit Kleinkindern (noch) keine Selbstverständlichkeit ist. Es gibt inzwischen einige Evidenz dafür, dass das mit Kindern im Alter ab neun Monaten regelmäßig der Fall ist und z.B. mit dem Zeigen einhergeht, was die Anforderung der Ko-Orientierung erfüllt.³⁵ Gut belegt ist weiterhin die Als-ob-Behandlung, die eindrücklich zeigt, dass und wie Erwachsene kindliches Verhalten so behandeln, als ob darin vollwertige Beiträge zur Interaktion zum Ausdruck kommen würden. Kompetentere Interaktionspartner:innen leisten also eine kommunikative Überbewertung des Verhaltens weniger kompetenter Partner:innen: »The mother from the start treats the child's cries as if they were intended to be more or less particular messages.«³⁶

Hier ist eine generelle Intentionalitätsunterstellung am Werk, die sich nicht nur in frühen vorsprachlichen »Interaktionsformaten« findet, sondern auch später in der Interaktion mit zweijährigen Kindern mit Bezug auf die Herstellung von Dialogen und noch später in der Interaktion mit Siebenjährigen mit Bezug auf Erzähl- und andere »Diskursfähigkeiten«.³⁷ Interaktionsfähigkeit erweist sich auch in diesem Kontext als eine flexible und interaktiv emergente Größe, die von der vorsprachlichen Ko-Orientierung und Ko-Ordination bis in anspruchsvolle Formen verbaler Ko-Operation (im Sinne von Dialog- und Diskursfähigkeiten) reicht. Speziell die Intentionalitätsunterstellung zeigt sehr anschaulich, wie die schon für das Anlaufen von Interaktion konstitutive Unterscheidung von Information und Mitteilung in und mit Interaktion hergestellt wird, obwohl sie von außen betrachtet das fragliche Verhalten systematisch überbewertet. Darin zeigt sich sehr anschaulich die Robustheit der Interaktion, was den Umgang mit womöglich problematischen Ressourcen betrifft; die Als-ob-Behandlung ist Ausdruck einer generellen wechselseitigen Anpassung der Beiträge der Interaktionsteilnehmer:innen an ihr Gegenüber, genauer gesagt: an ihr jeweiliges Fremdbild

von der Interaktionspartnerin oder dem -partner. Dieser Rezipientenzuschnitt der Interaktionsbeiträge ist in der Konversationsanalyse vielfach betont und belegt worden (»recipient design«). In der Erwachsenen-Kind-Interaktion, aber auch generell in Konstellationen, in denen die Interaktionsfähigkeit unter den Anwesenden asymmetrisch verteilt ist, zeigt sich die entwicklungsrelevante Kraft dieser Interaktionsdynamik. Sie erlaubt es als wichtiger Bestandteil eines systematischen »scaffolding«, dass der oder die weniger interaktionskompetente Beteiligte situationsflexibel in der »Zone der nächstfolgenden Entwicklung« agieren kann.³⁸

Asymmetrische Verteilungen von Interaktionsfähigkeit in der Interaktion zwischen Menschen zeigen sich auch im Fall spezieller Erkrankungen, Verletzungen oder anderer Beeinträchtigungen, die eine im Alltag routinemäßig unterstellte Interaktionsfähigkeit offenkundig massiv einschränken können. In diesem Zusammenhang sind u.a. Interaktionen mit Menschen mit Demenz in den letzten Jahren vermehrt in den Fokus der Forschung gerückt. Empirische Studien belegen, dass Interaktionsfähigkeit in diesen Interaktionen als flexible Größe situationsspezifisch manifestiert wird. Z.B. ergeben sich manifeste Einschränkungen mit Blick auf die symmetrische Beteiligung an der Bewältigung der Aufgaben der verbalen Interaktion. Die Interaktionspartner:innen werden z.B. mit Blick auf das gemeinsame Beitragen zu einem Thema (»Themaorganisation«) offenbar toleranter, was ihre Bereitschaft betrifft, ihre Ansprüche an thematische Kohärenz zurückzunehmen, während Beeinträchtigungen bei der Selbst- und Fremddarstellung schneller zu weiterreichenden Irritationen zu führen scheinen. Es wird berichtet, dass auch bis in spätere Phasen der Demenz eine »reine Form der körperlich-responsiven Interaktion« erhalten bleibt.³⁹ Die Beteiligung an der Situierung, speziell an der Ko-Orientierung und Ko-Ordination, gehört also anscheinend nicht nur zu den früh erworbenen und früh manifestierten Interaktionsfähigkeiten, sondern auch zu den noch lange erhaltenen und noch spät manifestierten Fähigkeiten. Es erstaunt nicht, dass offenbar auch in diesem Zusammenhang die Als-ob-Behandlung auftritt. Angehörige und Pflegende neigen dazu, Verhalten als Anzeichen für »größere kognitiv-mnemonische Fähigkeiten zu interpretieren, als tatsächlich noch vorhanden sind.«⁴⁰ Generell sind aus der Perspektive z.B. von Pflegenden im Umgang mit »Kommunikationslosen« Strategien der »simplifizierenden Zuordnung« in der Forschung gut belegt: »Im Stationsalltag werden Menschen ohne Kommunikation wie Menschen mit eingeschränkter Kommunikation behandelt, wenn das (quasi-)inverständlich auf sie bezogene und an ihnen ausgebrachte Handeln auf den

Zustandserhalt des körperlichen Wohls abzielt.«⁴¹ Diese Aussage bezieht sich auf Menschen mit schweren erworbenen Hirnschädigungen, die in der fraglichen Studie als »Kommunikationslose« bezeichnet werden. Interviews mit Pflegenden belegen, dass kleinste Zeichen für Zustimmung »intuitiv ›erahnt‹« werden, auch wenn die Unterstellung einer Mitteilungsabsicht problematisch bleibt. Das gilt auch, wenn z.B. durch musiktherapeutische Angebote körperliche Veränderungen bei den Patientinnen und Patienten ausgelöst werden sollen und diese Patientinnen und Patienten auf solche »Reaktionen« hin beobachtet werden. Wir haben es hier mit einem Übergangsbereich zu tun, in dem anders als in der Erwachsenen-Kind-Interaktion und offenbar auch in der Interaktion mit Menschen mit Demenz nicht auf einer mehr oder weniger gesicherten Grundlage von Ko-Orientierung und Ko-Ordination aufgebaut werden kann und in dem eher Informationen als Mitteilungen gesammelt werden. In diesem Kontext ist es besonders wichtig, die Frage der Interaktionsfähigkeit nicht auf der Grundlage der aus guten Gründen professionell oder emotional voreingenommenen Perspektive der Teilnehmenden zu beantworten, sondern einzelfallorientiert an den Erscheinungsformen des Verhaltens selbst festzumachen, um rekonstruieren zu können, ob und in welchem Ausmaß sich Spuren von interaktiv emergenter Interaktionsfähigkeit an der Oberfläche nachweisen lassen.⁴² Auch dabei können sich offensichtlich uneindeutige Situationen ergeben – bei denen es sich dann aber tatsächlich um Phänomene im Gegenstandsbereich fragwürdiger Kopräsenz handeln dürfte. Es ist also ein Unterschied, ob die Als-ob-Behandlung auf der Grundlage erfolgreicher Situierung aufbauen und ihre Wirkung erzielen kann oder als Strategie von Akteurinnen und Akteuren gewertet werden muss, um damit klar zu kommen, dass es gerade *keine* belastbare Evidenz für Wahrnehmungswahrnehmung und die Unterscheidung von Information und Mitteilung gibt und die »Bereitschaft zur Idealisierung« z.B. von »physiologischen Regungen« offenbar groß ist.⁴³ Gleichwohl zeigt diese Strategie – und zeigen insbesondere Interviews mit Angehörigen und Pflegenden –, dass und wie Alltagskonzepte von Interaktionsfähigkeit aussehen und dass die Unterstellbarkeit von Intentionalität offenbar zu solchen Alltagscollagen von Erwartungen an den Anderen als »Person« unverzichtbar dazugehört.⁴⁴

Weitere mit Blick auf den Personenstatus möglicher Interaktionsbeteiligter kritische Fälle ergeben sich beim Überschreiten der Gattungsgrenze und insbesondere mit der Zuwendung zu Formen und Funktionen des Kontaktes mit Haustieren (»interspecies communication«). Derartige Fälle werfen die Frage nach der Einzigartigkeit menschlicher Kommunikation und In-

teraktion auf, wie sie auch mit Blick auf das kommunikative Verhalten von Menschenaffen und nicht menschlichen Primaten gestellt und untersucht werden. Zugleich tun sich hier Möglichkeiten des Einblicks in phylogenetisch womöglich frühe(re) Formen basaler Kopräsenz auf. Schließlich gibt es seit einigen Jahren eine sehr lebhaft diskutierte Interaktionsfähigkeit von künstlicher Intelligenz, speziell in der Mensch-Maschine- bzw. Mensch-Roboter-Interaktion, die einen Blick in die Zukunft nicht länger exklusiv menschlicher Interaktion und technisch refigurierter Kopräsenz erlauben (»materialized refiguration«, »machine agency«).⁴⁵ So wie mit der Menschwerdung ko-evolutionär ein Wandel von Kopräsenz einhergeht, zeichnet sich womöglich auch mit technischen Innovationen im Bereich der künstlichen Intelligenz eine Refiguration von Kopräsenz ab, mit der die Grenzen dessen, was als Interaktionsfähigkeit in Anspruch genommen und an der Oberfläche der Interaktion manifest wird, weiter verschoben werden.

Erhellend sind in diesem Zusammenhang Situationen, in denen es Menschen mit »sozialen Robotern« zu tun bekommen, die als potentiell interaktionsfähige »humanoide« Akteure in Erscheinung treten, d.h. nicht mehr wie eine Kaffeemaschine oder ein Computer »bedient« und »benutzt« werden wollen, sondern ihrem Gegenüber dem Anspruch nach als eigenständiger Akteur gegenüberzutreten, der als Interaktionspartner adressierbar ist. Hier ist nicht der Ort, den Stand der technischen Entwicklung aufzuarbeiten, um zu entscheiden, ob die Technik diesem Anspruch bereits gerecht wird. Es geht nicht um eine Diagnose der Interaktionsfähigkeit von künstlich intelligenter Robotik.⁴⁶ Worum es geht, ist eine Einschätzung der Erscheinungsformen der »Begegnung« zwischen Mensch und Roboter im Hinblick auf die Frage, ob dabei so etwas wie Interaktionsfähigkeit manifest und greifbar wird. Gefragt sind also empirische Untersuchungen solcher Situationen anhand der Erscheinungsformen, in denen sich der Kontakt zwischen Mensch und Maschine manifestiert. Dass es sich dabei um Kommunikation handelt, ist vor dem Hintergrund des in dieser Schrift vertretenen Kommunikationsbegriffes unstrittig. Strittig ist, ob es sich dabei nach wie vor um Kommunikation unter der Bedingung von Benutzbarkeit und Bedienbarkeit handelt oder aber ob dabei neuerdings Kommunikation unter der Bedingung von Anwesenheit, d.h. als Interaktion, anläuft und Kopräsenz zwischen Mensch und Maschine entsteht. Wir haben einen solchen Fall eingangs bereits anhand des Anrufes eines Voicebots bei einem Dienstleistungsunternehmen besprochen und festgestellt, dass unter den Bedingungen der Telefonkommunikation und für den klar umrissenen Anlass einer Reservation Telekopräsenz offenbar erfolgreich hergestellt wer-

den konnte. Wie aber sieht es aus, wenn die Begegnung vor Ort und sozusagen »face-to-face« stattfindet? Wie vorliegende Daten zeigen, ist die Herstellung von Kopräsenz zwischen Mensch und Maschine viel anspruchsvoller und stör-anfälliger als die Herstellung von Telekopräsenz zwischen Mensch und Voice- oder Chatbot. Das hat in erster Linie mit der multimodalen Komplexität der Wahrnehmungswahrnehmung vor Ort zu tun. Wie wir oben gezeigt haben, ist die Interaktionseröffnung eine eminent anspruchsvolle Aufgabe, die mit dem Zusammenspiel von menschlichem Körper und natürlicher Sprache evolutionäre Errungenschaften besonderer Qualität in Anspruch nehmen kann. Damit nun unter den Bedingungen des Mensch-Maschine-Kontaktes Interaktion nahegelegt werden kann, muss die für Kopräsenz konstitutive Wahrnehmungswahrnehmung suggeriert werden. Am Telefon gelingt das mit der situativ angemessenen Begrüßung (s.o.). In einer Begegnung von Museumsbesucherinnen und -besuchern mit dem humanoiden Roboter Nadine im Paderborner Heinz-Nixdorf-Forum, die Florian Muhle für ein Forschungsprojekt aufgezeichnet hat,⁴⁷ gelingt die Suggestion von Wahrnehmungswahrnehmung dadurch, dass Nadine eine humanoide Gestalt aufweist, die auf einem Stuhl an einem Tisch mit Computer sitzt, mittels Kamera mit einer Sensorik ausgestattet ist, die Bewegungen im Raum registriert, und über eine Motorik verfügt, die es erlaubt, von einer »Grundposition« über eine »Begrüßungs-« in eine »Unterhaltungsposition« zu wechseln, in dem sie den Oberkörper und Kopf aufrichtet, die Blickrichtung verändert und mit dem Arm winkt, so dass der Anschein erzeugt wird, dass sie gesehen hat, dass sich jemand ihr annähert, verbunden mit der Suggestion, dass sie also gesehen haben könnte, dass sie gesehen wurde, so dass einer Kopräsenzsituation durch Körperbewegungen, Blickrichtung und Armbewegung Ausdruck verliehen wird. Wie Muhle in seiner Analyse zeigt, lässt sich eine solche Herstellung von Kopräsenz aber mit der Annäherung des Menschen an die auf ihrem Platz stationär verankerte Nadine nicht lange aufrechterhalten. Das hat auch damit zu tun, dass es mit zunehmender Annäherung vermehrte Hinweise darauf gibt, dass es sich bei Nadine um einen ausgestellten Roboter handelt und dass z.B. keine mündliche Begrüßung durch Nadine erfolgt. In dem analysierten Fall kommt es stattdessen zu einer längeren Schweigephase, in der nichts zwischen den Beteiligten passiert und die Besucherin oder der Besucher die auf dem Tisch vorhandenen Instruktionen studiert (»Sprechen Sie mit mir!«). Die Situation, die sich daraufhin ergibt, zeigt eine Bedien- und Benutzungssituation, die darin besteht, so zu tun, als ob ein Dialog mit Nadine möglich wäre (Frage einer Besucherin, laut, langsam und deutlich gesprochen: »Wie alt bist du?«,

Antwort Nadine: »Ich sehe aus wie dreißig, aber ich bin im Jahr 2015 hergestellt worden.«). Dadurch entsteht eine paradoxe Situation, weil sie die Bedienung der technischen »affordances« von Nadine in Form der Beteiligung an einem Dialog aus Frage und Antwort verlangt, bei dem Nadine als adressierbare Akteurin auftritt. Muhle zieht aus diesen und anderen Beobachtungen den Schluss, dass es sich weder um eine Interaktion handelt (in und mit der Nadine als Interaktionspartnerin inkludiert wird) noch um eine reine Bediensituation, bei der die Maschine als eigenständige Interaktionspartnerin exkludiert wird. Er spricht diesbezüglich in Anlehnung an das Konzept der Adressabilität von einer technischen Adresse. Tatsächlich zeigen auch andere Untersuchungen, dass eine in der Kontaktaufnahme womöglich augenblickshaft hergestellte Kopräsenz in der Regel nicht lange Bestand hat, sobald es Indizien dafür gibt, es beim Gegenüber mit einem Roboter zu tun zu haben. Die daraufhin gleichwohl zustande kommenden »Dialoge« zeigen, wie der menschliche Interaktionspartner oder die menschliche Interaktionspartnerin in dieser Situation die für ihn oder sie selbst offenbar relevanten und situativ erwartbar gemachten Konzepte von Interaktionsfähigkeit abfragt bzw. durch gezielte Initiativen (auch körperlicher Art) auszuprobieren und zu provozieren versucht.⁴⁸ Das lässt Dialoge entstehen, die immer wieder die Bedingungen der Möglichkeit von Interaktionsfähigkeit auf die Probe und auch in Frage stellen, aber damit doch gleichzeitig und nolens volens schon eine bestimmte Version von Interaktionsfähigkeit in Anspruch nehmen müssen. Dafür braucht es – anders als in der Erwachsenen-Kind-Interaktion – keine Als-ob-Behandlung und keine Ko-Orientierung mehr. Wenn man sich vorstellt, dass es sich in diesen Kontakten um klar umrissene aufgabenorientierte Dialoge handelt (wie beim Reservieren eines Tisches oder eines Termins), kann man sich auch vorstellen, dass die abstrakte Frage der Interaktionsfähigkeit hinter die Faktizität der erledigten konversationellen Aufgabe(n) mehr und mehr zurücktreten wird und womöglich neue Erscheinungsformen von Dialog- und Antwortpräsenz hervorbringen wird.⁴⁹ Kopräsenz und die mit ihr verbundene Reflexivität der Wahrnehmung könnten dabei in den Hintergrund treten, so dass es zu einem Hybrid verbaler Mensch-Maschine-Interaktion *ohne* Kopräsenz kommen könnte.⁵⁰ Hervorgegangen aus der Situierung mit Ko-Orientierung, Ko-Ordination und Ko-Operation könnte sich die verbale Interaktion von dieser Situierung in einer Kopräsenzsituation gleichsam freimachen und damit in gewisser Weise das wiederholen, was mit und durch Schrift schon einmal erfolgreich stattgefunden hat: die Entlastung und Emanzipation sprachlicher Kommunikation von der Bedingung der Kopräsenz. Vielleicht liegt die Poin-

te der Mensch-Maschine-Interaktion darin, dass sie diesen Schritt jetzt auch als Interaktion unter Anwesenden möglich macht. Und vielleicht ist es ja das, wofür wir soziale Roboter brauchen: für Interaktion *ohne* Kopräsenz.

Telekopräsenz: Anwesenheit aus der Ferne

Im Rahmen der Diskussion der Interaktionsfähigkeit von Robotern und künstlich intelligenter Technik wird häufig eine andere kommunikationstechnologische Entwicklung in einem Atemzug (mit-)thematisiert, bei der das »Gegenüber« zwar (häufig noch) ein Mensch, aber nur vermittelt durch Übertragungstechnik und Kommunikationstechnologien (wie Telefone oder Videokonferenztools) »anwesend« ist.⁵¹ Ich kann mein Gegenüber hören bzw. hören und sehen und auch hören (und sehen), dass ich gehört (und gesehen) werde, obwohl die oder der Andere räumlich weit entfernt von mir sein mag. Je nach disziplinärem Interesse, konzeptionellem Ausgangspunkt und gewählter Terminologie ergeben sich vielfältige Ein- und Zuordnungen solcher Kommunikationsformen in ein kaum noch überschaubares Spektrum von »Fernkommunikation«, »Medienkommunikation« und »technisch vermittelter Kommunikation«. Dabei werden unterschiedliche Kriterien zur Einheitenbildung und -abgrenzung genutzt, ohne dass wir diese Systematisierungsversuche an dieser Stelle im Einzelnen nachzeichnen wollen. In der konversationsanalytisch orientierten Literatur wird an dieser Stelle häufig und zumeist implizit auf den Gegensatz von direkter vs. vermittelter Interaktion zurückgegriffen. Er ist nicht unproblematisch, weil die Vorstellung einer unmittelbar-direkten Interaktion Teil des Kopräsenzmythos und der mit ihm verbundenen »Face-to-face«-Stilisierung ist.⁵² Wir schlagen deshalb einen anderen Weg ein und nehmen ausgehend vom eingeführten Begriff der Kopräsenz mit dem Konzept der Telekopräsenz eine konzeptionelle und terminologische Alternative auf, die an ausgewählte Vorarbeiten anschließen kann. Telekopräsenz meint eine Kommunikationsbedingung, bei der die für Kopräsenz konstitutive Reflexivität der Wahrnehmung der Beteiligten gegeben ist, aber nicht auf körpereigenen Kommunikationsorganen, sondern auf Kommunikationstechnologien zum gleichzeitigen Senden und Empfangen von Audio- oder/und Videodaten beruht. Neben dem Telefongespräch und der Videotelefonie sind Videokonferenzen (mittels Plattformen wie »Zoom«, »Teams«, »Jitsi« oder »Webex«) ein seit der COVID-19-Pandemie massenhaft verbreitetes Beispiel für Interaktionen auf der Grundlage von Te-

lekopräsenz. Damit erweist sich Kopräsenz entgegen vieler zeitgenössischer Diagnosen als eine robuste und wandlungsfähige Kommunikationsbedingung, die zur Aufgabe des mit dem Kopräsenzmythos verbundenen Raum-Zeit-Apriori und zur Relativierung der damit ebenfalls verbundenen Stilisierung von Unmittelbarkeit zwingt. Die Unterscheidung von Kopräsenz und Telekopräsenz macht darüber hinaus auf den mit Telefongesprächen und Videokonferenzen verbundenen Wandel aufmerksam. Er liegt darin, dass neue Kommunikationstechnologien nicht nur Alternativen zu Kopräsenz als Kommunikationsbedingung hervorbringen, die – wie Lesbarkeit – nicht mehr auf der Anwesenheit (= Wahrnehmungswahrnehmung) der Beteiligten, sondern auf ihrer kommunikativen Erreichbarkeit beruhen.⁵³ Neue Kommunikationstechnologien können auch die materiale Refiguration von Kopräsenz selbst vorantreiben (»material refiguration«) und Kopräsenz in einer bisher kaum gekannten Weise medial »technologisieren« (»technologized interaction«, »mediatization of face-to-face encounters«).⁵⁴ Wie wir zeigen werden, lässt sich diese Refiguration von Kopräsenz zu Telekopräsenz (system-)theoretisch als Wiedereintritt (»re-entry«) der grundlegenden Unterscheidung von Anwesenheit und Erreichbarkeit auf der einen Seite der Unterscheidung verstehen. Im Bereich der anwesenheitsbasierten Interaktion gibt es damit einen auf spezifische Weise markierten, d.h. unterscheidbaren Fall: den Fall der Telekopräsenz.

I.

Telefongespräche und Videoanrufe (verschiedenster Form) realisieren zweifellos Interaktion in dem von uns favorisierten engeren Sinn von reflexiver Wahrnehmung und Kopräsenz. Unter den Fachleuten für Interaktion besteht jedenfalls kein Zweifel daran, dass diese Formen technischer vermittelter Kommunikation zum Gegenstandsbereich der Interaktion gehören. Telefongespräche liefern nicht nur in der Anfangsphase der kalifornischen Konversationsanalyse das bevorzugte Datenmaterial, das maßgeblich dazu beigetragen hat, die für die verbale Interaktion so prägenden Aufgaben des Anfangens, des Sichabwechselns und des Abschließens zu explorieren. Sie belegen auch auf höchst anschauliche Weise, dass Wahrnehmungswahrnehmung nicht mit der physischen Anwesenheit der Beteiligten an einem Ort (Gleichörtlichkeit) gleichgesetzt werden darf oder, anders gesagt, (soziale) Anwesenheit auch unter (physisch) Abwesenden hergestellt werden kann, wenn und in dem Maße, in dem wahrgenommen werden kann, dass wahrgenommen wird – was mit dem Ab-

nehmen des Hörers nach dem Klingeln unterstellbar der Fall ist. Insbesondere die Analyse von Interaktionseröffnungen hat gezeigt, dass Kopräsenz eine Errungenschaft ist, die auch über Distanzen hinweg hergestellt werden kann. Entsprechend heißt es am Ende einer Zusammenfassung des einschlägigen Forschungsstandes zu »distant openings«: »In the light of these observations and findings, co-presence comes out as a practical and materially/technologically-supported multimodal accomplishment that can be achieved in different ways, in distant calls and virtual communication, but also while sharing the same location«.⁵⁵ Dieses Fazit schließt auch Videokonferenzen ausdrücklich mit ein, wie Lorenza Mondada zu Beginn einer Studie über »Eröffnungen und Prä-Eröffnungen« von Videokonferenzen deutlich macht: »Videokonferenzen haben mit Telefongesprächen gemeinsam, dass sie vermittelte Interaktionen zwischen Teilnehmern an verschiedenen Orten sind; der Eintritt in die Interaktion setzt voraus, dass über die technische Vorrichtung des Partners eine Verbindung hergestellt wird.«⁵⁶

Allerdings ging die frühe Konversationsanalyse von Telefongesprächen noch mit der Entdeckung des Gegenstandes der verbalen Interaktion überhaupt einher, so dass die Sonderbedingungen des Telefonierens eher eine Fußnote wert waren, als dass sie in eine tiefergehende Auseinandersetzung mit den medial-technologischen Besonderheiten im Vergleich zur »Face-to-face«-Interaktion eingemündet wären.⁵⁷ Im Gegensatz dazu haben neuere Entwicklungen speziell im Bereich der Videotelefonie von Anfang an die analytische Auseinandersetzung mit dem Einfluss der kommunikationstechnologisch-medialen Bedingungen provoziert, häufig vor dem Hintergrund dessen, was man über verbale Interaktion bereits aus der frühen Konversationsanalyse wusste. Dabei steht aber von Anfang an nicht die Frage im Mittelpunkt, ob es sich dabei (noch) um Interaktion handelt, sondern wie sich die technologischen Ressourcen auf die Bewältigung der typischen Aufgaben verbaler Interaktion auswirken und dafür nutzbar gemacht werden können.⁵⁸ Auch in diesem Sinn handelt es sich, so der Sprachgebrauch, um »vermittelte« Interaktionen. Dass wir es bei Telefon- und Videokommunikation grundsätzlich mit Interaktion und Kopräsenz zu tun haben, geht aus den Analysen also auf eher implizite Weise hervor. Es lässt sich aber an den Daten selbst relativ leicht nachweisen. Sofern die Wahrnehmung der Wahrnehmung der oder des Anderen nicht verlässlich (genug) durch das Sehen-des-Gesehen-Werdens auf dem Bildschirm und daran anschließende Begrüßungssequenzen gesichert scheint, erscheinen typischerweise situationsreflexive Fragen und Kommentare, die direkt auf die Wahrnehmbarkeit Bezug nehmen (Kannst

Du mich hören?)⁵⁹ und als Versuche der Vergewisserung über Kopräsenz verstanden werden können.⁶⁰ Noch und gerade am Umgang mit den viel beschriebenen kritischen Momenten der technologischen Fragmentierung von Kopräsenz zeigt sich so die Orientierung der Beteiligten an der Erwartung wechselseitiger Wahrnehmbarkeit.

Allerdings ist auch klar, dass es mit dieser zumal gemeinhin geteilten Feststellung nicht getan ist. Entsprechend zeigen die inzwischen vorliegenden Untersuchungen zu Videokonferenzen aus ethnomethodologisch-konversationsanalytischer Richtung, wie sich die technisch-mediale Vermittlung von Anwesenheit in den Erscheinungsformen der Kommunikation selbst empirisch manifestiert.⁶¹ Dazu gehört die schon angesprochene Vergewisserung wechselseitiger Wahrnehmung im Rahmen der Eröffnung (und »Prä-Eröffnung«), die es beim Telefonieren so (längst) nicht (mehr) gibt, solange nicht konkrete Irritationen auftauchen (weil der oder die Andere beim Zuhören z.B. länger auf ein Rückmeldesignal verzichtet hat). Die allgegenwärtige Vergewisserung wechselseitiger Wahrnehmung hat auch damit zu tun, dass Blickkontakt nicht möglich ist bzw. durch den Blick in die eigene Kamera lediglich suggeriert werden kann. Es handelt sich eben nicht um eine »Face-to-face«-, sondern um eine »Face-to-screen (with camera)«-Konfiguration, zumal zwischen Blick in die Kamera und Blick auf den Bildschirm entschieden werden muss (»eye-contact-dilemma«). Überdies stellen (mögliche) technische Probleme (»technological disruptions«) eine nicht nur bei der Eröffnung, sondern durchgängig relevante thematische Ressource dar, auf die bei der Themenorganisation zurückgegriffen werden kann und mit deren Ausnutzung zugleich Aspekte der Kommunikationstechnologie relevant gemacht werden (können). Schließlich stellt sich auch die Frage nach dem gemeinsamen Interaktionsraum neu. So ist bekannt, dass es mit den neben- und untereinander angeordneten kleinen Monitoren der Kameraaufzeichnungen der Anwesenden zu einer Vielzahl von sichtbaren Räumen in Form von Bildschirmhintergründen kommt. Der von den Anwesenden geteilte Wahrnehmungs-, Bewegungs- und Handlungsraum ist also nicht der jeweilige Raum der Beteiligten, in dem sie vor ihrem Bildschirm und der Kamera sitzen, sondern das auf der Videokonferenzplattform arrangierte Tableau mit übertragenen Bildern, auf denen die Einzelräume als Hintergrund der Köpfe ausschnitthaft sichtbar, aber nicht automatisch auch für die Interaktion relevant werden. Mit dem Übergang von Kopräsenz zu Telekopräsenz fällt aber nicht einfach der geteilte Interaktionsraum weg (»fractured« und »fragmented ecologies«),⁶² sondern es verändern und verlagern sich die Möglichkeiten seiner Herstellung. Körpergebundene Verfahren der Ko-Orien-

tierung, Ko-Ordination und Ko-Operation wie das Zeigen müssen neu adaptiert werden (»fractured indexicality«)⁶³ und rufen neue Verfahren des kameraorientierten Präsentierens hervor (»showing sequences«).⁶⁴ Auch ist daran zu denken, dass der gebaute und mit Möbeln gestaltete Raum als Interaktionsarchitektur unter Anwesenden eine starke Ressource darstellt, auf die die Beteiligten bei der Schaffung eines gemeinsamen Hiers zurückgreifen – etwa schon dann, wenn sie »nur« zusammen an einem Tisch in einem Zimmer Platz nehmen.⁶⁵ Diese Ressource, die weitgehend körperlich durch Verfahren der Navigation und Positionierung erschlossen wird, fällt unter telekopraesent Anwesenden weg. Die Herstellung des Interaktionsraums ist auf mehr als verkörperte Motorik und Sensorik angewiesen, eben auf Kameras und Mikrophone, häufig auch auf explizite sprachliche Instruktionen und Kommentare. An die Stelle gebauter und möblierter Interaktionsarchitektur(en) treten die »affordances«⁶⁶ der Videokonferenzplattformen, die Kopräsenz refigurieren. Markanter Ausdruck der mit Telekopraesenz einhergehenden Refiguration sind die »Gesichterwände« mit einem Tableau von Kacheln für die Beteiligten, sogenannte »talking heads configurations«.⁶⁷ Damit können dann auch Aspekte von privater Räumlichkeit (im Sinne des »Wohnens« und des »Zuhause«) in die kommunikative Zone ragen und erwartbar mitwahrgenommen und damit zum Interaktionsraum werden. Hier stehen wir wohl erst noch am Anfang einer Entwicklung, die das eigene Wohnen mehr und mehr telekopraesenzfähig umgestalten wird (Smarthome), so dass wir, um telekopraesent verfügbar sein zu können, z.B. nicht mehr darauf angewiesen sein werden, ein Endgerät mit Internetzugang auf- und anzustellen, auf dessen beschränkte Reichweite(n) wir angewiesen sind.

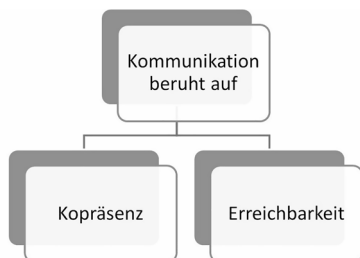
Die hier genannten Phänomene zeigen, dass und wie sich die zum Einsatz gebrachten Kommunikationstechnologien an der Oberfläche der Interaktion manifestieren und Kopräsenz im Sinne von Telekopraesenz refigurieren. Dass viele dieser Phänomene typisch sein dürften für den Umgang mit einer neuen Kommunikationstechnologie, kann nicht verwundern. Jedenfalls kann man mit Blick auf den Forschungsstand ohne Zweifel festhalten, dass man auch mit einem engen Interaktionsbegriff auf die Entwicklung neuer Kommunikationstechnologien gewinnbringend reagieren kann, also weder empirisch noch theoretisch immun ist gegenüber Evolutionen im Gegenstandsbereich. Die in Videoanrufen hergestellte Anwesenheit wollen wir aber nicht einfach als weiteren Fall von Kopräsenz verbuchen, den man mit Blick auf Abweichungen und Besonderheiten bei der Bearbeitung der bekannten Aufgaben verbaler Interaktion untersuchen kann (wie es bereits in Ansätzen erfolgt ist). Vielmehr

sehen wir darin einen besonderen Fall, der nicht nur empirisch, sondern auch konzeptionell genauer unter die Lupe zu nehmen ist, wenn man über den letztlich nicht befriedigenden Gegensatz von vermittelter und direkter Interaktion hinauskommen will. Wir müssen dazu zunächst kurz daran erinnern, dass es Alternativen zu Kopräsenz als Kommunikationsbedingung gibt.

II.

Mit Schrift gibt es schon seit langem eine wirkmächtige und einflussreiche Alternative zum Sichttreffen unter Anwesenden und also auch zur Interaktion. Diese Alternative zu Kopräsenz besteht in der Erreichbarkeit von Anderen als Lesenden, sei es eines Briefes oder einer Zeitung oder eines Buches. Neben Kopräsenz als Kommunikationsbedingung ist damit so etwas wie Erreichbarkeit, konkreter gesagt, die »Lesbarkeit« (des Textes) als Kommunikationsbedingung getreten.⁶⁸ Die Entwicklung der COVID-19-Pandemie hat uns jüngst vor Augen geführt, wie weitgehend die auf Erreichbarkeit und Lesbarkeit basierenden Formen der Kommunikation Interaktion unter Anwesenden ersetzen können. Daneben ist deutlich geworden, dass und wie Kommunikationstechnologien nicht nur die Ersetzung, sondern auch die Refiguration von Kopräsenz als Kommunikationsbedingung vorantreiben können. In diesem Kontext ist besonders die Videokonferenz in Erscheinung getreten, mit der wie schon illustriert ein Wandel hin zu Telekopräsenz stattgefunden hat. Wie ist dieser Wandel theoretisch einzuordnen?⁶⁹ Wir gehen dazu im ersten Schritt von unserer Ausgangsunterscheidung zwischen Anwesenheit und Erreichbarkeit aus und unterscheiden damit grundsätzlich zwei Fälle von Kommunikation (s. Abb. 19).

Abb. 19: Kopräsenz vs. Erreichbarkeit (eigene Darstellung)



Stark vereinfacht und verkürzend gesagt unterscheiden wir damit das Sichunterhalten in Gesprächen vom Schreiben-und-Lesen von Texten. Entscheidend für diese Unterscheidung ist allerdings nicht, wie das Beispiel suggerieren mag, der Unterschied zwischen gesprochener und geschriebener Sprache oder der zwischen Nähe und Ferne der Beteiligten. Entscheidend ist, dass mit der Etablierung einer von Kopräsenz unabhängigen kommunikativen Erreichbarkeit eine Zäsur eintritt zwischen Sendung und Empfang (Produktion und Rezeption) der Erscheinungsformen der Kommunikation. Diese Zäsur hat in der Regel einen zeitlichen Aspekt, insofern an die Stelle der Gleichzeitigkeit das Nacheinander von Sendung und Empfang tritt. Damit sind die Erscheinungsformen der Kommunikation nicht länger das Dokument einer faktisch vollzogenen Kommunikation in Form einer Interaktionsepisode mit Anfang und Ende, sondern das Dokument einer möglich und wahrscheinlich gemachten Kommunikation (z.B. in Form der Lektüre), in und mit der nicht länger Kopräsenz, sondern Lesbarkeit bzw. allgemeiner gesagt Erreichbarkeit hergestellt wird.

Die Unterscheidung ist also nicht trivial, und sie versteht sich auch nicht von selbst. Soziologisch hat sie mit dem Hintergrund der Luhmann'schen Systemtheorie einen anderen Resonanzraum als in der Linguistik, in der sie gewissermaßen ihre stillen Begleiterinnen hat, die nicht genau das Gleiche unterscheiden, aber oft in ähnlichem Zusammenhang auftauchen. Das betrifft die eben schon angedeuteten Unterscheidungen von Mündlichkeit und Schriftlichkeit, Nähe und Ferne und Synchronizität vs. Asynchronizität (auf die wir gleich noch zurückkommen). Was den soziologischen Hintergrund betrifft, schließen wir an eine Begrifflichkeit an, die Niklas Luhmann im Zuge seiner These von der »Unwahrscheinlichkeit der Kommunikation« verwendet hat. Wir haben diese These schon an anderer Stelle aufgegriffen und erläutert⁷⁰ und können also gleich zur Sache kommen. Eine spezifische Unwahrscheinlichkeit der Kommunikation ergibt sich, wenn und in dem Maße, in dem Kommunikation sich über den Kreis jeweils Anwesender hinaus verbreiten soll: Diese Unwahrscheinlichkeit

bezieht sich auf das *Erreichen* von Empfängern. Es ist unwahrscheinlich, dass eine Kommunikation mehr Personen erreicht, als in einer konkreten Situation anwesend sind. Das Problem liegt in der räumlichen und zeitlichen Extension. Das Interaktionssystem der jeweils Anwesenden garantiert in praktisch ausreichendem Maße Aufmerksamkeit für Kommunikation, und es zerbricht, wenn man erkennbar kommuniziert, dass man nicht kommunizie-

ren will. [...] Selbst wenn die Kommunikation bewegliche und zeitbeständige Träger findet, wird es daher unwahrscheinlich, dass sie Aufmerksamkeit voraussetzen kann. In anderen Situationen haben die Leute etwas anderes zu tun.⁷¹

Erreichbarkeit (jenseits von Anwesenheit) ist also das Problem, um das es hier geht. Damit ist nicht nur gemeint, dass die Kommunikation dafür auf materiale Substrate und eine ausgebauten Infrastruktur von Übermittlung und Zustellung angewiesen ist, im Falle von Sprache also auf so etwas wie Schrift und so etwas wie Post – das auch. Allein damit ist aber nicht schon sichergestellt, dass Andere sich auf diese Kommunikation überhaupt einlassen (wollen), wenn die Hydraulik der Kopräsens wegfällt. Erreichbarkeit impliziert also auch, dass die Kommunikation attraktiv genug wird, um potentielle Empfänger:innen anzusprechen und auf sich aufmerksam zu machen. Wie auch immer man dieses Unwahrscheinlichkeitstheorem weiter ausbuchstabieren mag: Erreichbarkeit versteht sich bei Luhmann als alternative Kommunikationsbedingung zu Anwesenheit/Kopräsens. Sie ist Schwelle und Errungenschaft zugleich, wenn sich Kommunikation über den Kreis Anwesender hinaus ausbreiten soll, die an einer Kommunikation Beteiligten also nicht schon über Anwesenheit und Wahrnehmungswahrnehmung faktisch erreicht werden, sondern auch auf andere Weise erreichbar (gemacht) werden müssen, so dass möglich und wahrscheinlich wird, dass auch Abwesende miteinander kommunizieren können und ihre Inklusion durch Verbreitungsmedien erwartbar (gemacht) werden kann. Dazu haben die Entwicklung und Verbreitung von Schrift mit den Errungenschaften der Alphabetschrift und des Buchdrucks maßgeblich beigetragen.⁷²

Als Alternative zu Anwesenheit hat der Begriff der Erreichbarkeit bei Luhmann auch noch an anderer Stelle eine zentrale Bedeutung. Der Begriff findet sich auch bei der Unterscheidung von Interaktion, Organisation und Gesellschaft als jeweils eigenständigen Möglichkeiten der Konstitution sozialer Systeme. Auch diesen Kontext haben wir bereits thematisiert, so dass wir hier daran erinnern können, dass jedem dieser Systemtypen von Kommunikation eine eigene Kommunikationsbedingung zugrunde liegt: Im Fall von Interaktion ist das eben Anwesenheit (im Sinn von Kopräsens), im Fall von Organisation Mitgliedschaft und im Fall von Gesellschaft als dem umfassenden Sozialsystem Erreichbarkeit. Da Gesellschaft nach Luhmann gegenwärtig nur als »Weltgesellschaft« begriffen werden kann, geht es hier letztlich um weltweite kommunikative Erreichbarkeit, die ohne die Errungenschaften der

Verbreitungsmedien nicht möglich wäre: »Soziologen und Historiker betonen in seltener Einmütigkeit, dass erst die kommunikationstechnologischen Revolutionen, von der Telegraphie bis hin zum Internet, die Voraussetzung für weltweite Vernetzungs- und Globalisierungsprozesse geschaffen haben.«⁷³ Bettina Heintz schließt daraus, dass Telekommunikation vielfach an die Stelle von Interaktion getreten ist und der Eindruck entstanden sei, dass Erreichbarkeit Anwesenheit mehr und mehr ersetzt und in der Bedeutung relativiert habe.⁷⁴ Es ist aber noch die Frage, ob man Erreichbarkeit als Konstitutionsprinzip von (Welt-)Gesellschaft umstandslos auf das Phänomen der Telekommunikation reduzieren kann. Zunächst wäre anzumerken, dass Telekommunikation im Sinn technisch vermittelter Kommunikation Interaktion nach unserem Verständnis nicht ausschließt und als Kategorie von daher problematisch ist – Telefongespräche realisieren ohne Zweifel *sowohl* Telekommunikation *als auch* Interaktion. Wie wir gezeigt haben, gilt das auch für Videokonferenzen, ohne dass man dafür den engen Begriff von Interaktion bzw. Kopräsenz »sensu« Goffman und Luhmann aufgeben muss. Dazu kommt, dass die Erreichbarkeit der Gesellschaft Anwesenheit und Kopräsenz einschließt und dass es Erreichbarkeit als Konstitutionsprinzip nicht erst seit den o.g. »kommunikationstechnologischen Revolutionen« gibt. Natürlich hat die Entwicklung der Kommunikationstechnologien dazu beigetragen, Erreichbarkeit als Alternative zu Kopräsenz zu profilieren. Aber es geht um mehr als nur unterschiedliche Formen von Kommunikation (wie es der implizite Gegensatz von Nah- und Fernkommunikation suggeriert): Mit der Bedingung der Erreichbarkeit wird Kommunikation von diskreten Interaktionsepisoden mit Anfang und Ende in die prinzipiell »zeitlose« Möglichkeit und Wahrscheinlichkeit kommunikativer Anschlussmöglichkeiten verschoben. Genau dafür steht das Suffix »-bar« in »Erreichbarkeit«. Im Gegensatz dazu steht Kopräsenz dafür, dass die Möglichkeit der Wahrnehmbarkeit der oder des Anderen unmittelbar umschlägt in die konkrete Faktizität der Wahrnehmungswahrnehmung. Deshalb zerfällt die anwesenheitsbasierte Kommunikation in Interaktionssepisoden mit Anfang und Ende, während die auf Erreichbarkeit basierende Kommunikation im Prinzip ihr Zustandekommen, und damit auch Anfang und Ende, mit eigenen (Bord-)Mitteln nicht (mehr) sicherstellen kann. Sie bleibt zwar wie jede Kommunikation

zeitgebunden, wird aber auf eine spezifische Weise unabgeschlossen. Am Beispiel von Lesbarkeit haben wir diesen abstrakten Gedanken an anderer Stelle fruchtbar gemacht.⁷⁵ An diesen soziologischen Begriff von Erreichbarkeit, der Lesbarkeit als eine prominente Realisierungsform einschließt, wollen wir mit der o. eingeführten Unterscheidung anknüpfen.

III.

Was den *linguistischen* Resonanzraum betrifft, in den die Ausgangsunterscheidung von Anwesenheit und Erreichbarkeit fällt, finden sich an vergleichbarer Stelle in der Regel Unterscheidungen, die ihre Kriterien aus dem Bezug auf die Medialität im Sinne der Art des Kontaktes, die Sprache, den Raum oder die Zeit gewinnen. Sie sind alle bereits angeklungen. Aus Gründen der Veranschaulichung haben wir »Face-to-face«-Interaktion (*Kopräsenz*), Textkommunikation (*Lesbarkeit*) sowie Telefon- und Videogespräche (*Telekopräsenz*) jeweils zugeordnet (s. Abb. 20 und 21).

Abb. 20a und 20b: Kontakt: direkt vs. vermittelt & Sprache: mündlich vs. schriftlich (eigene Darstellung)

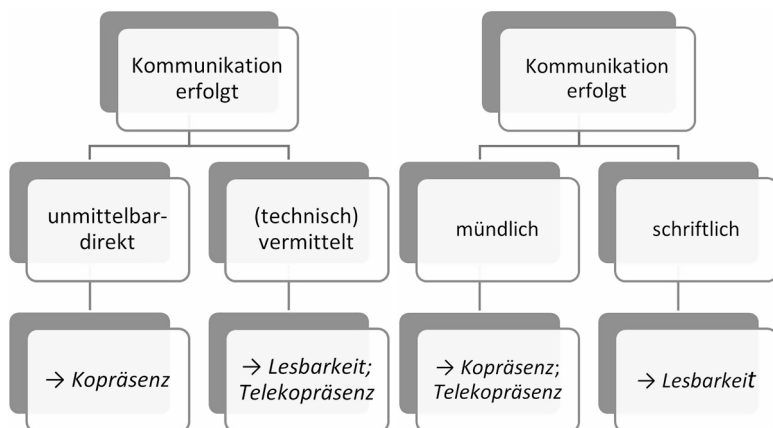
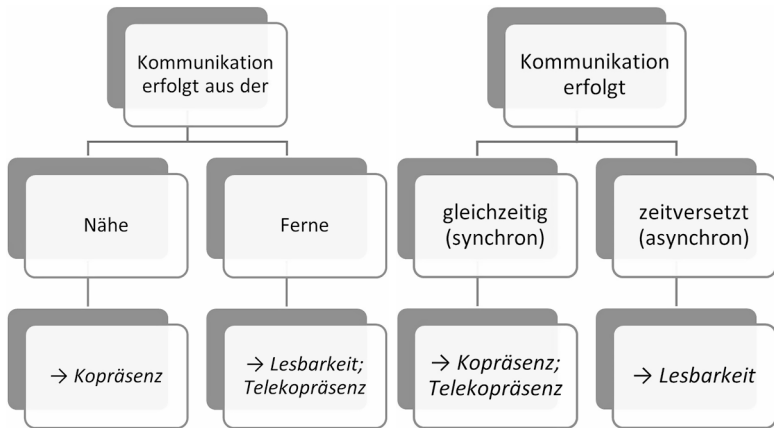


Abb. 21a und 21b: Raum: Nähe vs. Ferne & Zeit: synchron vs. asynchron (eigene Darstellung)



Die Unterscheidung von unmittelbar-direkt vs. technisch bzw. medial vermittelt hat offensichtlich mit dem Einsatz von Kommunikationstechnologien zu tun, die von Mündlichkeit und Schriftlichkeit offenkundig mit Sprache und Sprachlichkeit, die von Nähe und Ferne offensichtlich mit Raum und Räumlichkeit und die von Gleichzeitigkeit und Ungleichzeitigkeit offenkundig mit Zeit und Zeitlichkeit. Keine dieser Unterscheidungen fällt mit der von uns eingeführten Unterscheidung von Anwesenheit und Erreichbarkeit zusammen, obwohl sie in vergleichbarem Zusammenhang immer wieder genannt werden und – wie wir schon gesehen haben – bis in die Terminologie hinein ausstrahlen (»technisch vermittelte Kommunikation«, »Tele-« oder »Fernkommunikation«). Hinzu kommt, dass die (einfachen) Fälle von Kopräsenz und Lesbarkeit anhand dieser Unterscheidungen jeweils gleich zugeordnet werden. Der Fall der Telefon- und Videokommunikation (Telekopräsenz) führt dann aber schon zu unterschiedlichen Zuordnungen. Die Problematik des Gegensatzes von direkt vs. vermittelt mit seiner Nähe zum Mythos Kopräsenz haben wir bereits angedeutet. Es kommt hinzu, dass es auch so etwas wie Telekopräsenz gibt, die nach unserer Auffassung trotz des Einsatzes von Kommunikationstechnologien auf die Seite der Kopräsenz gehört. Die gleiche Problematik zeigt sich beim Gegensatz von (räumlicher) Nähe vs. Ferne. In der Form von Telekopräsenz ist Nähe gewissermaßen auch in der Ferne realisierbar, insofern Kommunikationstechnologien die Reflexivität von Wahrnehmungen auch unabhängig von Gleichörtlichkeit möglich machen.⁷⁶ Bleiben die sprach- und zeitbasier-

ten Unterscheidungen, die beide zu gleichen Zuordnungen führen, was aber einer genaueren Analyse nicht standhält. So ist der Gegensatz von Mündlichkeit und Schriftlichkeit, der historisch womöglich mit dem von Anwesenheit und Erreichbarkeit zusammengefallen ist, inzwischen auf beiden Seiten der Unterscheidung nicht mehr verlässlich; auch gesprochene Sprache kann aufgrund der Möglichkeit ihrer Konservierung und Übertragung unabhängig von Kopräsenz eingesetzt werden (z.B. auf Anrufbeantwortern oder in Rundfunk und Fernsehen). Umgekehrt kann Schrift vereinzelt auch in Kopräzenssituationen auftreten (etwa beim Nutzen der Chatfunktion während gleichzeitiger Videokonferenz). Die an der Erscheinungsform von Sprache orientierte Unterscheidung ist also längst nicht mehr trennscharf (was zuweilen für begriffliche Verwirrungen sorgt), und spätestens mit der »sekundären Oralität«⁷⁷ fällt die Mündlichkeit nicht mehr mit dem Nähe- und Gleichzeitigkeitspol zusammen.

Bleibt die an der Zeitstruktur orientierte Unterscheidung von Gleichzeitigkeit und Nacheinander, Synchronizität und Asynchronizität; diese Unterscheidung kommt der von Anwesenheit und Erreichbarkeit vielleicht am nächsten.⁷⁸ Sie erlaubt eben auch die Differenzierung zwischen Technologien, die synchrone Kommunikation (in »Echtzeit«) ermöglichen und solchen, die »asynchrone Kommunikation« (mit Zäsur zwischen Sendung und Empfang) implizieren, auch wenn das Schreiben-und-Lesen von Text in elektronischen Umgebungen auf tragbaren Kommunikationsgeräten oftmals schon zu einer quasisynchronen Verständigung geführt hat.⁷⁹ Wir sehen die Synchronizität aber als eine Begleiterscheinung der Reflexivität der Wahrnehmung an, mit der Sender:innen immer auch Empfänger:innen und Empfänger:innen immer auch Sender:innen sind. In einer Situation des Nacheinanders von Sendung und Empfang ist diese Konfundierung der Variablen ausgeschlossen. Umgekehrt führt aber Synchronizität nicht zwangsläufig schon zur Wahrnehmungswahrnehmung, wie das Beispiel der Live-Übertragung einer Fernsehsendung (in Echtzeit) belegt. Wir sehen die Zeitstruktur deshalb als nachgeordnet an. Das wird im Übrigen auch dadurch bestätigt, dass der Dualismus von Sendung und Empfang (kommunikativer Erscheinungsformen), auf dem die Unterscheidung von synchron vs. asynchron beruht, ein nachträglicher ist, der erst mit und nach der Erfindung von Schrift und weiteren Verbreitungsmedien denkmöglich geworden ist und deshalb an die Kopräzenssituation von außen herangetragen wird und ihr letztlich nicht gerecht werden kann.⁸⁰

Auch ohne dass wir hier auf die mit den genannten Unterscheidungen verbundenen Forschungstraditionen näher eingehen können,⁸¹ zeigt sich, wie

ausgehend von den genannten Unterscheidungen die Weichen unterschiedlich gestellt werden; wer primär Nah- von Fernkommunikation unterscheidet, wird das Telefongespräch (nomen est omen) und die Videokonferenz wie selbstverständlich zur Fernkommunikation zählen; wer synchrone von asynchroner Kommunikation unterscheidet, wird das Telefongespräch und die Videokonferenz gleichwohl und zweifelsfrei zur synchronen und schließlich auch zur mündlichen Kommunikation zählen. Telefongespräch und Videokonferenz bilden zwar jeweils zusammen ein Gespann, das nicht getrennt voneinander auftaucht; aber nicht immer taucht dieses Gespann auf der gleichen Seite der Unterscheidungen auf. Wir sehen darin ein oberflächliches Indiz dafür, dass hier Klärungsbedarf besteht, wenn man über deskriptive Terminologie und oberflächliche Begriffsbestimmungen hinauskommen will. Zugleich sollte transparent(er) sein, warum wir mit der Differenz von Anwesenheit und Erreichbarkeit eine Ausgangsunterscheidung gewählt haben, die ohne Referenzen auf Kommunikationstechnologien, Raum, Zeit und Sprache auskommt. Kopräsenz steht zwar der Einheit von Raum/Ort und Zeit nahe (zur gleichen Zeit am gleichen Ort), ist damit aber als Kommunikationsbedingung nicht angemessen erfasst. Wenn und in dem Maße, in dem wahrgenommen werden kann, dass (die eigene Wahrnehmung der oder des Anderen von der oder dem Anderen auch) wahrgenommen werden kann, entsteht Kopräsenz. Ein solcher zugegeben abstrakter Begriff hat den Vorteil, dass er offenlässt, *wie* diese Wahrnehmungswahrnehmung im konkreten Einzelfall bewerkstelligt oder eben auch ausgeschlossen werden kann (und wie sie sich evolutionär entwickeln kann). Zwei Leute im gleichen Zimmer, die einander z.B. aufgrund einer Trennscheibe weder akustisch noch visuell wahrnehmen können (oder nur einseitig wahrnehmen können), sind also im hier angesetzten Sinn nicht kopräsent. Zwei Leute, die über ein Telefon sicherstellen können, dass sie sich wechselseitig hören können, sind aber sehr wohl kopräsent (auch wenn zwischen ihnen der Atlantik liegt). Gleichwohl wollen wir damit nicht die Besonderheit der Telefonkommunikation gegenüber einem Gespräch mit Gleichörtlichkeit außer Acht lassen. Wenn es nicht der Gegensatz von Nähe und Ferne ist, was ist dann die entscheidende Besonderheit?

IV.

Bleiben wir zunächst der Einfachheit halber beim Telefongespräch. Auch wenn man es wie hier vorgeschlagen auf der Seite der Kopräsenz verortet (wegen der

Realisierung von Wahrnehmungswahrnehmung), wird man nicht übersehen wollen, dass in diesem Fall mehr im Spiel ist:

Der Eintritt in die Interaktion setzt voraus, dass über die technische Vorrichtung des Partners eine Verbindung hergestellt wird. [...] [D]ie Herstellung der Verbindung selbst (besteht) aus einer ganzen Reihe von Aktivitäten – Abheben des Hörers durch den Anrufer, Wählen, Freizeichen (bzw. auf Seiten des Angerufenen Klingeln), eine mehr oder weniger lange Folge von Klingeltönen und schließlich Abnehmen des Hörers.⁸²

Erst mit dem Abnehmen des Hörers nach dem Klingelzeichen beginnt die Interaktion, weil vorher weder auf Seiten der oder des Anrufenden noch der oder des Angerufenen wahrgenommen werden kann, dass wahrgenommen wird. Die oder der Anrufende muss bis zum Abnehmen des Hörers der oder des Angerufenen davon ausgehen, dass die oder der Andere nicht »da« ist oder womöglich aus Gründen gerade nicht abnimmt, obwohl er »da« ist. Auch die Anruferin bzw. der Anrufer könnte es sich anders überlegen und seinerseits vor dem Abnehmen auflegen. Die oder der Andere könnte dann wohl wissen, dass jemand (und heutzutage oft auch: wer) gerade seine Nummer gewählt hat, aber die Möglichkeit der Wahrnehmungswahrnehmung bestand nicht. Deshalb bleiben sowohl das Klingeln wie auch das Nichtabnehmen des Hörers interaktiv folgenlos. Erst mit dem Abnehmen des Hörers nach dem Klingelton kann wahrgenommen werden, dass wahrgenommen wird. Wenn kein Freizeichen ertönt (weil die oder der Andere schon wieder aufgelegt hat), kann schon das Ausbleiben des Freizeichens und die wie kurz auch immer ausfallende Stille als Beleg dafür gehört werden, dass auf der jeweils anderen Seite jemand »da« und »dran« ist. Wer als Erste bzw. Erster zu sprechen anfängt, macht dann schon den zweiten Zug. Insofern ist es berechtigt zu sagen, dass mit dem Abnehmen des Hörers gehört werden kann, dass (zu-)gehört wird – auch wenn für einen kurzen Augenblick noch gar nichts zu hören ist. Man sieht daran zweierlei: dass und wie das Klingeln retrospektiv durch das Abnehmen z.T. der Interaktion gemacht wird (und gemacht werden muss); und dass die Beteiligten anstelle wahrnehmbarer empirischer Evidenz aufgrund ihrer Vertrautheit mit der Praxis des »Telefonierens« routinemäßig unterstellen können, dass der oder die Andere mit dem Abnehmen des Hörers »dran« ist. Es muss daher niemand etwas sagen wie: »Ich bin dran«, damit es losgehen kann. Mit dem Sprechen verdichtet sich dann nicht anders als in der »Face-to-face«-Situation der Verdacht der Wahrnehmungswahrnehmung zur Gewissheit, und es ist

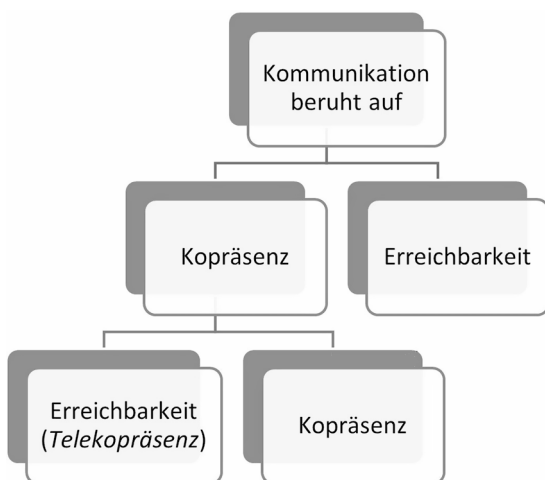
bekannt, dass diese Gewissheit zu schwinden beginnt, wenn von der jeweils anderen Seite länger nichts zu hören ist. Wir haben diesen »Vorlauf« der Interaktion in der Darstellung bewusst etwas in die Länge gezogen, um zu illustrieren, dass beim Telefongespräch die Herstellung von Kopräsenz durch reflexive Wahrnehmung »technologisiert« ist. Das gilt für den Vorlauf, für den es auf beiden Seiten einen »Apparat« braucht, und es gilt für die Eröffnung, für die es eine »Verbindung« braucht. Man kann also mit Fug und Recht behaupten, dass in diesem Fall Erreichbarkeit im Spiel ist: Ich muss die Andere oder den Anderen offenkundig auch *erreichen* können, bevor wir einander in die Zone wechselseitiger Wahrnehmbarkeit geraten (können).

Dieser technisch-mediale Vorlauf und die auf Erreichbarkeit beruhende Herstellung reflexiver Wahrnehmung unterscheiden die Telefoninteraktion von Interaktionen, die nur auf der Grundlage verkörperter sinnlicher Wahrnehmung (»naked senses«)⁸³ zustande kommen, so dass Erreichbarkeit als Kommunikationsbedingung unmittelbar umschlägt in Wahrnehmungswahrnehmung. Wenn wir diese Unterscheidung im Bereich der kopräsenzbasiererten Interaktion in den Termini der Ausgangsunterscheidung selbst umsetzen wollen, ohne dabei doch wieder auf die anderen genannten Parameter zurückzugreifen, bietet es sich an, die Unterscheidung von Kopräsenz und Erreichbarkeit auf sich selbst zu beziehen und auf der einen Seite der Unterscheidung einen besonderen Fall als »markierten Fall« einzuführen, so dass es einen Wiedereintritt der Unterscheidung von Kopräsenz und Erreichbarkeit auf der Seite der Kopräsenz gibt (s. Abb. 22).

Luhmann hat diese Form des Wiedereintritts von Unterscheidungen in dem von ihnen Unterschiedenen (in Anlehnung an George Spencer Brown) als Denkfigur des »re-entry« populär gemacht.⁸⁴ Wir können damit im Bereich der kopräsenzbasiererten Interaktion einen markierten Fall unterscheiden, bei dem Kopräsenz ihrerseits auf Erreichbarkeit beruht, wie wir das am Beispiel des Telefongesprächs illustriert haben: Wahrnehmungswahrnehmung setzt einen technologisch-mediatisierten Vorlauf voraus und wird in einer auf Erreichbarkeit basierenden Weise hergestellt. Wir haben damit einen Spezialfall hybrider Kommunikation vor uns, bei dem sich zwei gegenüberstehende Merkmale gleichsam kreuzen. Am Fall des Telefonierens kann man sich gut klarmachen, was dabei passiert: Die Interaktion wird anfälliger für Störungen, weil sie sich nicht mehr auf die Bordmittel der beteiligten Körper verlassen kann, sondern auf Medientechnologien (z.B. in Form der Übertragung, der Sendung und des Empfangs elektrischer Signale) und Medieninstitutionen (z.B. eine Art »Fernmeldewesen«) angewiesen ist. Wie unsere Skizze zur

empirischen Erforschung von Videokonferenzen zeigt (s.o. I.), manifestiert sich diese Abhängigkeit von voraussetzungsreicher Medientechnologie auf der Oberfläche der Erscheinungsformen selbst. Die situationsreflexiven Kommentare zur Vergewisserung über Kopräsenz sind das beste Beispiel dafür.

Abb. 22: Wiedereintritt der Unterscheidung von Kopräsenz und Erreichbarkeit auf der Seite der Kopräsenz (eigene Darstellung)



Trotz dieser Abhängigkeit von Kommunikationstechnologie bleiben das Telefongespräch und die Videokonferenz aber unzweifelhaft Interaktion; die Übertragung ersetzt ja nicht die Gleichzeitigkeit der Wahrnehmungswahrnehmung zugunsten einer Aufsplittung in die Sendung und den Empfang von Signalen: Wer spricht, hört gleichzeitig zu, und wer zuhört, spricht auch dann, wenn sie oder er schweigt, ist also gleichzeitig Sender:in und Empfänger:in. Am Telefon ist das unmittelbar daran ersichtlich, dass der, der spricht, darauf hört, ob er auch gehört wird, und dass die oder der, die oder der zuhört, auch (gelegentlich) spricht, damit ihr oder sein Zuhören auch gehört werden kann. Wie die Beispiele aus dem seit der COVID-19-Pandemie sprunghaft angestiegenen Umgang mit Videokonferenztools zeigen, bilden sich vergleichbare Routinen hier erst noch aus. Dabei kommt es auch zu

Brüchen: Z.B. führt das Ausschalten der Kamera und des Mikrophons der Beteiligten (sofern sie nicht gerade »dran« sind), nicht selten dazu, dass die Gewissheit von Wahrnehmungswahrnehmung brüchig wird, weil tatsächlich nicht mehr gesehen werden kann, dass zugesehen (und zugehört) wird und die aktuelle Sprecherin oder der aktuelle Sprecher nicht mehr gleichzeitig auch Zuhörer:in ist, so dass sich eine Art schleichender Übergang von Kopräsenz zu Erreichbarkeit einspielt (der von den Beteiligten nicht selten auch kommentiert wird). Ein derartiger Übergang ist nur möglich, wenn Kopräsenz mit Erreichbarkeit konfundiert ist. Ohne diese Konfundierung ist mit dem Ende der Wahrnehmungswahrnehmung und der von ihr abhängenden Interaktionsepisode auch das Ende der Kommunikation gegeben. Um diese Besonderheit von Telefon- und Videokommunikation begrifflich festzuhalten, wollen wir, wie schon angekündigt, von Telekopräsenz sprechen. Wir nehmen damit einen Vorschlag von Shanyang Zhao auf, der eine Taxonomie verschiedener Fälle von Kopräsenz vorgeschlagen hat.⁸⁵ Mit Telekopräsenz bezeichnet er Fälle, in denen die Anwesenheit des Anderen nicht auf körperlicher, sondern auf elektronisch herbeigeführter Nähe beruht:

Physical proximity is an area within which naked human senses can reach, and electronic proximity is an area within which electronically extended human senses can reach. Through the mediation of a telephone or a Personal Digital Assistant (PDA), for example, individuals not mutually present in the same locale can reach each other at a distance in real time.⁸⁶

Abhängig davon, ob die oder der Andere selbst (»in person«) oder in Form einer Nachbildung (»through simulation«) anwesend ist, unterscheidet er zusätzlich noch zwischen »corporeal telecopresence« und »virtual telecopresence«.⁸⁷ Wir nehmen diese Taxonomie hier auf, weil sie im Gegensatz zum ebenfalls eingeführten Begriff von Telepräsenz⁸⁸ keinen Zweifel daran lässt, dass es um Telekopräsenz geht.

Einen Eindruck vom Wandel, der mit der Etablierung und Verbreitung von Telekopräsenzsituationen einhergeht, hat die COVID-19-Pandemie vermittelt; es ist genau dieser Typus hybrider Kommunikation, der sich seit dem Ausbruch der COVID-19-Pandemie ebenso rasant wie der Virus selbst verbreitet hat. Kopräsenz ist mit der Pandemie also nicht nur erschwert, verhindert und ersetzt, sondern auch *weiterentwickelt* worden. Sie ist vermehrt und verstärkt auch dann möglich gemacht worden, wenn die Interaktionspartner:innen zur gleichen Zeit an unterschiedlichen Orten und dort jeweils für sich allein sind. Man den-

ke dazu eben vor allem an all die inzwischen schon stark verbreiteten Softwaretools für »Meetinglösungen« in »kristallklarer HD-Qualität«. So ist »Zoom« nicht nur »the leader in modern enterprise video communications« (so die vollmundig werbende und maßlos übertreibende Selbstdarstellung), sondern zur Bezeichnung für eben einen neuen Typus von Kommunikation geworden. Unter dem Stichwort »zoom fatigue« werden bereits erste Begleiterscheinungen dieser Kommunikation kritisch diskutiert.⁸⁹ Diese neuen Formen von Telekopräsenz lassen es – wie vor Jahren schon einmal mit der flächendeckenden Verbreitung des Telefons in jedem Haushalt – vermehrt erlebbar werden, dass und wie Kopräsenz übertragungstechnisch-medial hergestellt werden kann. Anders als beim Tippen und Lesen von Mitteilungen und Nachrichten kann ich durch die Übertragung von Stimme und Bild nicht nur die Andere oder den Anderen wahrnehmen, sondern zumindest punktuell auch wahrnehmen, dass ich gerade wahrgenommen werde. Wir haben es hier also damit zu tun, dass eine abstrakte Hypothese – die Herstellung von Anwesenheit durch technische Hilfsmittel – zur Alltagserfahrung wird (wie schon einmal beim Telefon). Allerdings kommt es dabei, wie wir mit Rückgriff auf vorliegende empirische Untersuchungen illustriert haben, zu spezifischen Besonderheiten: Es ist, salopp formuliert, eben doch etwas anderes, ob man der oder dem Anderen direkt in die Augen oder auf einen Bildschirm bzw. in die Kameralinse des Laptops schaut.

Refigurationen von Kopräsenz durch Kommunikationstechnologie

Der Blick auf neuere Entwicklungen, die einen möglichen Wandel von Kopräsenz andeuten, weist darauf hin, dass und wie Refigurationen von Kopräsenz durch den Einsatz neuer Kommunikationstechnologien vorangetrieben werden. Das kommt als Fazit nicht überraschend und wird hier nicht zum ersten Mal kommentiert.⁹⁰ In der Sprache dieses Essays betrifft es die Ressourcen multimodaler Interaktion: Neue Informations- und Kommunikationstechniken, wie sie etwa bei der Sensorik und Motorik sozialer Robotik, künstlicher Intelligenz und Videogesprächen zum Einsatz kommen, treten mit ihren »affordances« als Ressource verbaler Interaktion prominent hervor. Wenn man die beiden Entwicklungen zusammennimmt, die wir in diesem Kapitel besprochen haben, die Technisierung von Interaktionsfähigkeit und die Verbreitung von Telekopräsenz, geht damit eine Verlagerung und Modifizierung der Bedeutung anderer Ressourcen der Interaktion einher. Das betrifft vor

allem die (Interaktions-)Architektur und mit ihr den Körper und die Sprache: Die architektonischen Benutzbarkeitshinweise, die als Navigations-, Interpretations- und Partizipationshinweise humanspezifische Fähigkeiten gleichzeitig fordern und fördern, werden durch kommunikationstechnische Benutzbarkeitshinweise verdrängt, deren Einfluss auf Körper und Sprache noch nicht absehbar ist. Einen wie immer eingeschränkten Vorgesmack auf zukünftige Entwicklungen hat diesbezüglich die COVID-19-Pandemie vermittelt. Die mit ihr einhergehende Konfiguration von Kopräsenz zu Telekopräsenz ist ein eindrückliches Beispiel dafür, dass und wie man sich einen Wandel der Kommunikationsbedingung der Kopräsenz vorstellen muss, was etwa die Herstellung des Interaktionsraums und mit ihr die Aufgabe der Situierung betrifft (s.u. I.). Dagegen zeigen andere Beispiele für institutionell ausdifferenzierte Interaktionsarchitekturen, wie Architekturen als Ressourcen Kopräsenz konfiguriert und dabei zugleich die Ersetzung von Architektur durch Kommunikationstechnik vorbereitet haben, mit der dann eine Verdrängung von Kopräsenz durch Erreichbarkeit einherzugehen scheint. Die Interaktionsarchitektur des Verkaufsschalters ist dafür ein sprechendes Beispiel (s.u. II.).⁹¹

I.

Einer der anhaltenden sozialen Effekte der COVID-19-Pandemie ist die Verbreitung von dem, was wir hier mit Telekopräsenz bezeichnet und abstrakt als eine Form der »Kreuzung« von Anwesenheit und Erreichbarkeit eingeführt haben. Allgemein gesprochen haben wir es hier mit einem aktuellen Beispiel für die Refiguration von Kopräsenz durch Kommunikationstechnik zu tun, bei der die Entwicklung nicht zum Ersatz von Kopräsenz durch Erreichbarkeit führt, sondern zu einer genuinen Weiterentwicklung von kopräsenzbasierter Interaktion.⁹² Ihre erste Welle hatte diese Refiguration von Kopräsenz zu Telekopräsenz bereits mit der Einführung und Verbreitung des Telefons erreicht. Viel spricht dafür, dass wir gerade – befeuert durch die Pandemie – eine zweite Welle der Verbreitung von Telekopräsenz mit der Nutzung des Internets und mobiler Endgeräte für Video- bzw. Webkonferenzen und Video- bzw. Bildtelefonie erlebt haben, auch wenn die dafür notwendige Technologie schon viel länger entwickelt und auch verbreitet worden ist. Was aber ist das Neue und wo liegen die möglicherweise auch längerfristig relevanten Veränderungen, die mit diesem neuerlichen Refigurationsschub einhergehen? Vieles von dem, was als »künstlich«, belastend oder erschwerend mit Blick auf Telekopräsenz wäh-

rend der Pandemie ins Feld geführt worden ist, lebt offenkundig von der Hypostasierung eines Ur- und Reinzustandes »natürlicher« Kopräsens, den es nie gegeben hat und der dem hier beschriebenen Mythos Kopräsens entspricht. Ausgehend von diesem Mythos lässt sich dann etwa kritisch feststellen, dass in Videokonferenzen die Wahrnehmung eingeschränkt und zugleich stark fokussiert, Blickkontakt nicht möglich und der Sprecherwechsel erschwert seien (wie man z.B. in einem Artikel im Zürcher Tages-Anzeiger lesen konnte).⁹³ Das alles ist nicht ganz falsch, aber doch sehr oberflächlich. Auch im Gespräch unter unmittelbar Anwesenden ragt ja nicht alles automatisch in die kommunikative Zone, was für die Beteiligten sinnlich wahrnehmbar ist, sondern nur ein kleiner Ausschnitt wird in seiner Wahrnehmbarkeit systematisch wahrnehmbar gemacht – sonst bräuchte es die Anforderung der *Situierung* nicht (s. noch u.). Der Blickkontakt ist seit jeher sehr viel stärker regelbasiert und kontrolliert, als uns das bewusst ist – wie man heute mit Verfahren des mobilen Eye-Tracking auch gut dokumentieren kann.⁹⁴ Und das, was wir schließlich über den Sprecherwechsel aus der Konversationsanalyse wissen, stammt bekanntlich originär aus der Beschäftigung und Konfrontation mit Telefongesprächen, also aus der Beschäftigung mit Telekopräsens. Insofern braucht es zunächst den Hinweis, dass die auf Kopräsens beruhende Interaktion seit jeher kein Wohlfühlsetting ist, sondern ein durch und durch strukturiertes, regel- und gesetzmäßig ablaufendes Geschehen, das an die Beteiligten systematische Anforderungen stellt, darunter auch eine nicht zu unterschätzende Körperdisziplin. Wir haben deshalb betont, dass Anwesenheit grundsätzlich eine kommunikative Hervorbringung ist, die nicht einfach gegeben und da ist, sondern hergestellt wird. Man musste, vereinfacht gesagt, schon immer nicht nur anwesend sein, sondern auch *zeigen*, dass man gerade anwesend ist. Weil das so ist, ist Interaktion wandel- und entwickelbar, d.h. mit dem Entstehen neuartiger Technologien auch über den Kreis des Hier und Jetzt ausdehnbar, zugleich aber auch viel robuster und anpassungsfähiger an interaktionsfeindliche Umwelten, in denen z.B. die menschliche Sinneswahrnehmung massiv eingeschränkt ist. So ist Interaktion selbstverständlich auch unter Anwesenden möglich, die weder hören noch sehen können (und dafür ganz andere Sensoren für die Interaktion zugänglich machen). Abstrakter formuliert: Interaktion ist eine robuste Sozialform, die sich den Bedingungen, unter denen sie zustande kommen kann, nicht nur anpasst, sondern diese Bedingungen von Fall zu Fall auch reproduzierend (mit-)gestaltet und (weiter-)entwickelt. Das betrifft auch den Interaktionsraum.

Mit dem Übergang von Kopräsenz zu Telekopräsenz fallen nicht einfach der unter Anwesenden geteilte Interaktionsraum und damit die Anforderung der Situierung weg, sondern es verändern und verlagern sich die Möglichkeiten der Herstellung des Interaktionsraums und damit die Ressourcen für Ko-Orientierung, Ko-Ordination und Ko-Operation. Beispielsweise fällt die Interaktionsarchitektur, die weitgehend körperlich durch Verfahren der Navigation und Positionierung erschlossen wird (also mit humanspezifischer Motorik und Sensorik geleistet wird), unter telekopräsent Anwesenden als Ressource weitgehend aus. Die Herstellung des Interaktionsraums und mit ihr die Aufgabe der Situierung sind dann auf mehr als verkörperte Motorik und Sensorik angewiesen, eben z.B. auf Übertragungs- und Übermittlungstechniken in Form von Kameras und Mikrofonen, häufig auch auf explizite sprachliche Instruktionen und Kommentare. An die Stelle gebauter und möblierter Interaktionsarchitektur(en) treten die »affordances« der Videokonferenzplattformen. Damit werden dann auch Aspekte privater Räumlichkeit als Teil des Interaktionsraums erwartbar mitwahrgenommen, so dass das eigene Zuhause unter dem Druck der Interaktion allmählich als Smarthome zu einem Multimediahub umgestaltet werden dürfte,⁹⁵ mit dem die Grenzen zwischen Kopräsenz und Telekopräsenz dann immer mehr verwischen. Bis es so weit ist, werden wir allerdings wohl noch eine Weile damit leben müssen, uns wechselseitig darauf hinzuweisen, dass das Mikrophon nicht angeschaltet, das Kamerabild eingefroren oder die Stimme zu leise ist. Aber klar ist auch, dass wir es hier mit Übergangsphänomenen zu tun haben, wohingegen sich deutlich abzeichnet, dass die auf Telekopräsenz beruhende Kommunikation mit und durch Sprache so schnell nicht aus unserem Alltag verschwinden wird. Ob und wie das neben der Kommunikation und mit ihr auch die Sprache selbst verändern wird, ist im Moment noch nicht abzusehen, muss uns aber auch nicht weiter kümmern: Die Sprache ist als Medium nicht weniger robust und anpassungsfähig als die Interaktion, in der sie ihr soziales Zuhause hat.

II.

Die Refiguration von Kopräsenz durch Kommunikationstechnik betrifft nicht nur wie im Falle der Telekopräsenz den Wandel von Kopräsenz, sondern auch die Ersetzung von Kopräsenz durch Erreichbarkeit. Architektonische Benutzungshinweise können sich dabei als Schrittmacher für die Ermöglichung von Kommunikation ohne Kopräsenz erweisen. Die Interaktionsarchitektur des Fahrkartenschalters am Bahnhof ist dafür ein aufschlussreiches Beispiel,

auf das wir abschließend eingehen wollen. Man kann daran sehen und zeigen, dass die Ausdifferenzierung einer spezifischen Interaktionsarchitektur auf die aufkommende Nutzung des Bahnhofs und den dabei erforderlichen Kauf von Tickets vor Ort reagiert, d.h. einen spezifischen Typ von Kopräsens neu konfiguriert und damit möglich und wahrscheinlich macht. Es ist also ein Fall von Konfiguration durch Architektur, der sich in den Navigations-, Interpretations- und Partizipationshinweisen des Schalters als Ressource der Verkaufsinteraktion manifestiert. Man kann an diesem Fall weiter sehen und zeigen, dass diese spezifische Konfiguration von Kopräsens rund 150 Jahre später in den Verkauf von Tickets an einem Fahrkartenautomaten übergegangen ist. Der Schalter ist dabei nicht nur historisch, sondern auch interaktionstheoretisch der Vorläufer und Wegbereiter des Ticketautomaten, mit der sich der Verkauf von Fahrkarten von der Kommunikation via Kopräsens zugunsten der Bedienung und Benutzung eines Automaten frei gemacht hat. Interaktionsarchitektur trägt also zunächst als Ressource zur Konfiguration von Kopräsens bei – bevor auf dieser Grundlage später Kopräsens als Kommunikationsbedingung durch die Benutzbarkeit bzw. Bedienbarkeit einer Maschine und damit Interaktion zwischen Anwesenden durch Kommunikation qua Technik ersetzt werden kann.⁹⁶

Die Entwicklung der Architektur des auch heute noch auf vielen Bahnhöfen anzutreffenden Fahrkartenschalters, die bis in die zweite Hälfte des 19. Jahrhunderts reicht, hat ohne Zweifel maßgeblich zur modernen Konfiguration der Verkaufsinteraktion beigetragen. Trotz einer Vielzahl von Publikationen zur Geschichte der Bahnhofsarchitektur ist die Geschichte des Fahrkartenschalters allerdings noch nicht geschrieben. Verschiedene Quellen lassen darauf schließen, dass der Übergang vom frei stehenden Verkaufsraum oder »Kassenhäuschen« (auch »Billet-Haus« genannt, zunächst außerhalb des Bahnhofsgebäudes, aber auch früh als Teil der Diensträume) zu einer in der Empfangshalle befindlichen »Wand mit kleinen Durchgabe- und Sprechöffnungen« etwa Mitte des 19. Jahrhunderts weitgehend vollzogen worden ist.⁹⁷ Die Ersetzung der massiven Wand durch eine Glasfront und entsprechende Glasschalter mit eigenen Vorrichtungen zum Transfer und zum Durchlass von Geld und Tickets hat dann wohl vor allem zwischen 1860 und 1890 stattgefunden, einer Phase der »stilistischen Verfeinerung und Ausdifferenzierung der Bahnhofsbauten«.⁹⁸ Das Ergebnis dieses Prozesses ist der uns auch außerhalb des Bahnhofs mehr oder weniger geläufige »Schalter«, wie er bis vor Kurzem z.B. (noch) in der Schalterhalle des Hauptbahnhofs Zürich anzutreffen war (s.u. Abb. 23a bis 23c). Zu seinen charakteristischen Wiedererkennungsmerk-

malen gehört ein spezifisches räumliches Arrangement, das für die, die sich am Schalter kopräsent finden, einen mehr oder weniger abgeschlossenen Binnenraum mit zwei möglichen Positionen definiert: Man ist entweder *vor* oder *hinter* dem Schalter und damit einhergehend *innerhalb* oder *außerhalb* des Schalterraums.⁹⁹

Abb. 23a, 23b und 23c: Schalterhalle HB Zürich (Quelle: Hausendorf und Mondada 2017)



Diese räumlichen sind zugleich soziale Positionen: Hinter dem Schalter und innerhalb des Schalterraums befindet sich ein komplexer Arbeitsplatz, der u. a. mit einer Kasse und weiteren mehr oder weniger technisch anspruchsvollen Gerätschaften ausgestattet ist. Vor dem Schalter und außerhalb des Schalterraums ist ein Platz zum Ankommen und Verweilen definiert, zu dem u. a. eine Ablage- und Anlehn- bzw. Abstützfläche gehört. An dieser Stelle kommt etwas sehr Wichtiges hinzu: Die vom Schalter definierte Grenze zwischen innen und außen erweist sich unter spezifischen Bedingungen als durchlässig und durchdringlich. Es sind dies die »Durchgabe- und Sprechöffnungen« in der Wand, von denen oben bereits die Rede war, die im Verlauf der Ausgestaltung als verglaste Front in Form spezieller Durchreich- und Austauschvorrichtungen architektonisch ausdifferenziert werden. Am Hauptbahnhof Zürich ist das eine mit einem speziellen Hebel von innen bedienbare Drehscheibe (s. Abb. 24).

Abb. 24: Drehscheibe mit Hebel als Durchreich- und Austauschvorrichtung (Quelle: Hausendorf und Mondada 2017)



Mit Vorrichtungen wie diesen erweist sich die durch den Schalter definierte Grenze als Kontaktzone zwischen dem Innen einer spezifischen Institution und dem Außen derjenigen, die sich an die Institution wenden, um irgend etwas zu erhalten, das an dieser Stelle (unter bestimmten Bedingungen)

durchgereicht (»transferiert«) werden kann. Der Schalter definiert den Ort, an dem dieser Kontakt stattfindet und an dem die Institution auf ihre Kundschaft trifft. Das kann am Bahnhof passieren, aber ebenso gut am Zirkuswagen, auf einem Schiff oder im Botschaftskonsulat.¹⁰⁰

Die Ausgestaltung der Front vor dem Schalter durch eine abgegrenzte Scheibe, einen Ablagebereich und eine Durchgabe- und Sprechöffnung lässt sich ohne weiteres als architektonischer Navigationshinweis verstehen, der dem und der, die den Kontakt suchen, zeigt, wo der Platz zum Ankommen und zur Eröffnung des Anliegens ist. Zugleich handelt es sich um wiedererkennbare Elemente einer als Schalter verstandenen Architektur, die Navigationshinweise sind also auch Interpretationshinweise zum Wiedererkennen des Schalters (wie er z.B. namentlich in dem Hinweis auftaucht, dass »[d]er Schalter geschlossen [ist]«).¹⁰¹ Für die, die mit der spezifischen Praxis vertraut sind (z.B. der Praxis des Fahrkarten- oder Eintrittskartenverkaufs), handelt es sich speziell bei der Durchgabeöffnung, die auf Abb. 24 als Drehscheibe ausgeformt ist, weitergehend um Partizipationshinweise: Wer vor dem Schalter ankommt und dort verweilt, macht sich als Kundin oder Kunde der Institution mit dem Anliegen erkennbar, etwas erwerben zu wollen. Diese in der Schalterarchitektur materialisierten Navigations-, Interpretations- und Partizipationshinweise mögen trivial erscheinen. Als wichtige Ressourcen insbesondere der Situierung (Wo sind wir?), der Positionierung (Wer bin ich?) und der Rahmung (Was geht hier vor?) sind sie aber alles andere als trivial, weil sie erklären helfen, dass und warum die Verkaufsinteraktion am Schalter unmittelbar nach der Interaktionseröffnung mit einem Anliegen gestartet werden kann, was die thematische Organisation (Was kommt als Nächstes?) und den Sprecherwechsel (Wer kommt als Nächste oder Nächster?) enorm entlastet. Die Szene, die wir o. mit den Abb. 23a bis 23c eingefangen haben, stammt aus einem Schaltergespräch, das wie folgt beginnt:

SBB-Mitarbeiter (hinter dem Schalter). *Grüezi. Was hätten Sie gern?*

Kundin (vor dem Schalter). *Es Halbs Contenschwil Retour.*

SBB-Mitarbeiter. *Jawohl* (beginnt damit, Informationen in das Programm einzugeben)¹⁰²

Der Austausch enthält keinerlei sprachliche Spuren der Vergewisserung über die Situierung, Rahmung oder Positionierung der Beteiligten, weil bereits die Ankunft am Schalter diesbezüglich für Klarheit sorgt und eine ebenso effektive wie standardisierte Abarbeitung des Anliegens (in diesem Fall: des Kaufs

eines Tickets) ermöglicht. Die Drehscheibe zum Austausch von Ware gegen Geld manifestiert gleichermaßen die Erwartung eines Anliegens, das mit diesem Austausch erfüllt werden kann. Insofern erweist sie sich als sehr effektive Ressource für die Beendigung der Interaktion: Die Bedienung der Drehscheibe durch den Mitarbeiter leitet den Transfer ein, mit dem die Beendigung eröffnet werden kann.¹⁰³ Natürlich geht es am Schalter auch anders. Keine Interaktionsarchitektur kann determinieren, was sich unter denen, die kopräsent sind, ereignet. Aber sie kann das, was sich ereignen soll, möglich, wahrscheinlich und hochgradig erwartbar machen. Auf diese Weise trägt sie zu einem effektiven aufgabenorientierten Geschehen bei, das sich z.B. in einem sprachlich vorgeformten und entsprechend formelhaften Verlauf der verbalen Interaktion niederschlägt. In der gerade wiedergegebenen Eröffnung eines Schaltergesprächs zeigt sich das sehr anschaulich schon in der Formulierung des Anliegens (»Es Halbs Gontenschwil Retour«).

Es kommt somit zu einer immer stärkeren Standardisierung der Schalterinteraktion, die in ihrem Ablauf und den dafür notwendigen »Slots« vorhersehbar und planbar wird – was sich in den von uns aufgezeichneten Gesprächen z.B. in der auf dem Mitarbeitercomputer installierten und auf dem Bildschirm manifestierten Software zum Ticketverkauf zum Ausdruck kommt. Wir sehen hier eine Entwicklung am Werk, mit der die Interaktion am Schalter gleichsam domestiziert wird im Sinne klar definierter, institutionell »erlaubter« Anliegen. Die Ausdifferenzierung einer wiedererkennbaren Schalterarchitektur (wie sie in der zweiten Hälfte des 19. Jahrhunderts stattgefunden hat) ist insofern der materiale Ausdruck einer Standardisierung der an diesem Ort möglichen und erwartbaren Interaktion. Diese Standardisierung führt ihrerseits auf der Seite der Interaktionsbeteiligten und der am Schalter Anwesenden (»Agentinnen« bzw. »Agenten« und »Klientinnen« bzw. »Klienten«) zu einer zunehmenden »Habitualisierung«, mit der »Handlungen ohne Bewusstsein quasi-automatisch ablaufen«.¹⁰⁴ Zugespitzt gesagt: Am Schalter werden die Beteiligten zunehmend dazu gebracht, so zu agieren, »dass Maschinen ihr Handeln nachahmen können«.¹⁰⁵ Aus der Gegenwartsperspektive erweist sich der Schalter damit als Vorläufer der modernen, elektronisch gesteuerten Fahrkartenautomaten, wie sie seit etwa Mitte der 1990er Jahre auf Bahnhöfen anzutreffen sind: Mit seiner Glasfront mit Transfervorrichtungen erscheint der Schalter im Rückblick bereits wie ein Automat, der sich (nur) dadurch auszeichnet, dass darin (noch) ein Mensch (aus Fleisch und Blut) sitzt.¹⁰⁶ Die mit dieser Architektur konfigurierte Standardisierung der Verkaufsinteraktion und die Habitualisierung der damit verbundenen Handlungen haben

den Ablauf der Interaktion nicht nur hochgradig erwartbar, sondern gewissermaßen auch »programmierbar« gemacht. An die Stelle einer auf Kopräsenz beruhenden Interaktion zwischen Menschen vor und hinter dem Schalter ist die Benutzung bzw. Bedienung einer Maschine getreten. Insofern ist die Konfiguration von Verkaufsinteraktionen durch den Fahrkartenschalter ein Beleg dafür, dass Kopräsenz als Kommunikationsbedingung nicht nur wandlungsfähig ist, sondern mit der Weiterentwicklung ihrer Ressourcen auch als Wegbereiterin alternativer Kommunikationsbedingungen in Erscheinung treten kann. Unzweifelhaft handelt es sich bei der Bedienung eines Fahrkartenautomaten um Kommunikation. Aber sie beruht genauso unzweifelhaft nicht länger auf Kopräsenz (mit einer Person aus Fleisch und Blut im Innern des Automaten),¹⁰⁷ sondern auf Lesbarkeit und auf der Benutzbarkeit von Technik als alternativen Kommunikationsbedingungen, die Kommunikation von Kopräsenz entlasten. Wer vor einem Fahrkartenautomaten steht und anfängt, sich auf der Benutzeroberfläche zu orientieren, rechnet nicht mit einem Menschen hinter der Front des Automaten, sondern mit einer durch Technologie gesteuerten Maschine. In dem Maße, in dem solche Maschinen z.B. durch den Einsatz künstlicher Intelligenz »interaktionsfähig« werden, könnte es allerdings sein, dass sich diese Wahrnehmung verschiebt – und auch der Verkauf von Tickets am Automaten wieder als Interaktion refiguriert wird. Mit diesem Ausblick in eine ungewisse Zukunft wollen wir unseren Rück- und Ausblick auf Kopräsenz im Wandel beschließen.