

Ko-Konstruktion in der Mensch-Roboter-Interaktion

Kontingenz, Erwartungen und Routinen in der Eröffnung

KAROLA PITSCH

Contingency – interactional contingency – is not a blemish on the smooth surface of discourse, or of talk-in-interaction more generally. It is endemic to it. It is its glory. It is what allows talk-in-interaction the flexibility and the robustness to serve as the enabling mechanism for the institutions of social life.

E.A SCHEGLOFF (1996: 22)

1. EINLEITUNG

Will man technische Systeme so ausstatten, dass sie vom Nutzer mit den Mitteln natürlichsprachlicher Kommunikation bedient werden können, dann besteht eine zentrale Herausforderung darin, dass Mensch und Roboter mit grundlegend verschiedenen Annahmen über Handlungsstrukturen operieren (vgl. Suchman 1987): Während technische Systeme weitgehend auf Handlungslogiken basieren, die Planbarkeit von Abläufen voraussetzen, ist menschliche Kommunikation situiert und kontingent. Ihr Ablauf ist grundsätzlich offen, die einzelnen Schritte sind nicht im voraus planbar, sondern entwickeln sich lokal Zug um Zug als eine gemeinsame Leistung aller Interaktionsbeteiligten (*»interactional achievement«*), so dass das Herstellen von interaktiven Anschlüssen eine praktische Aufgabe für die Beteiligten darstellt. Zu ihrer Bearbeitung verfügen menschliche Kommunikationsteilnehmer über ein Set an Ressourcen, wie z.B. projektive Strukturen unterschiedlicher Reichweite (Streeck/Jordan 2009, Auer 2009, Dausendschön-Gay/Krafft 2009) mittels derer Erwartungen aufgebaut werden (Levinson 2006). Wenngleich jüngere Modellierungsansätze für Dialogsysteme beginnen, planungsbasierte, strukturalistische und probabilistische Ansätze miteinander zu verbinden (vgl. Lison 2015), ist bislang noch weitgehend ungeklärt, inwiefern technische Systeme zum Umgang mit der Situietheit und Kontingenz menschlichen Kommunikationsverhaltens befähigt werden können. Interaktionsanalytisch basierte Studien zur sog. Mensch-Roboter-Interaktion (HRI) sind daher zum einen interessant für die Weiterentwicklung von technischen Systemen, und zum an-

deren ermöglichen sie – durch die Diskrepanz der divergierenden Handlungsstrukturen und daraus entstehenden Situationen von »Krisenexperimenten« (Garfinkel 1967) – Einblicke in die Bedingungen von *interactional achievement*. Insbesondere erlaubt eine interaktionistische Perspektive, Mensch und Roboter als *ein* Interaktionssystem zu verstehen, das gemeinsam die anstehenden Interaktionsaufgaben bearbeitet (Pitsch et al. 2013). Das von Gülich/Dausendschön-Gay/Krafft (i.d.B.) vorgeschlagene Konzept der Ko-Konstruktion legt dabei den analytischen Fokus auf die Beiträge der Teilnehmer und ihre *accounts* für die gemeinsame Arbeit.

Vor diesem Hintergrund stellt der »Einstieg in eine Interaktion« einen interessanten Untersuchungsgegenstand dar, da er ein hohes Maß an Koordinierungsleistung und kleinschrittigen Entscheidungen über interaktive Anschlüsse erfordert. Wenn – wie in der von uns durchgeführten Studie (Pitsch/Wrede 2014)¹ – reguläre Besucher einer Kunstausstellung bei ihrem Rundgang zum ersten Mal auf einen (humanoiden) Roboter-Guide treffen, dann stehen sie vor der Aufgabe herauszufinden, über welche Kompetenzen das Gegenüber verfügt, worauf es reagiert, wie man es ggf. bedienen kann und welche Aktivitäten relevante Anschlusshandlungen im »Interaktionssystem Mensch-Roboter« darstellen. Mensch und Roboter müssen also nicht nur in Kontakt miteinander treten, sondern auch wechselseitig Zugänglichkeit herstellen. Während dieses in menschlicher Interaktion als ein kleinschrittiger multimodal-koordinativer Prozess erfolgt, in dem mindestens Verfügbarkeit, Wechselseitigkeit, Identifikation/Kategorisierung und die Interaktionsmodalitäten für eine mögliche bevorstehenden Interaktion etabliert werden müssen, findet es in der Mensch-Maschine/Roboter-Interaktion unter veränderten Bedingungen statt. In bisherigen Systemen wird der Einstieg in ein *focused encounter* zumeist als ein vordefinierter verbaler Handlungsablauf konzipiert (z.B. Lohse 2009, Schegloff 1986), so dass eine Interaktionsaufgabe – wie die Eröffnung – durch das vom Roboter vorgegebene Handlungsskript als beendet definiert wird und nicht – wie in menschlicher Kommunikation üblich – als Resultat einer situativen Herstellungsleistung und wechselseitigen Aushandlung darüber, wann die Bedingungen für den Übergang zur nächsten Aufgabe hergestellt sind (siehe Dausendschön-Gay/Gülich/Krafft i.d.B.).

Angesichts dieser Situation haben wir in der o.g. Studie den traditionellen verbal-sprachlichen Eröffnungsroutinen des Roboters aus menschlicher Interaktion inspirierte responsiv-dynamische Praktiken des *pre-beginning* (Signalisieren von Aufmerksamkeit und Wahrnehmung) vorangestellt. Im folgenden Text werden wir auf dieser Basis folgender Frage nachgehen: Wie organisieren Mensch und Roboter unter diesen Bedingungen gemeinsam den Einstieg in ein *focused encounter*? Welche Interaktionskompetenzen schreiben die Besucher dem Roboter zu und welche Erwartungen erwachsen daraus hinsichtlich relevanter Anschlussmöglichkeiten im Interaktionssystem? – Die Analyse wird drei Fallbeispiele mit unterschiedlichen interaktiven Voraussetzungen berücksichtigen. Implikationen werden hinsichtlich des Verständ-

1 Dieser Text geht aus den Projekten »Interaktion & Raum« (Volkswagenstiftung) und »Interactional Coordination and Incrementality in Human-Robot-Interaction« (CITEC, EXC 277) hervor. Besonderer Dank gilt Sebastian Wrede für die Kooperation an der Robotik-Schnittstelle.

nisses von Ko-Konstruktion im ›Interaktionssystem Mensch-Roboter‹ und der Anforderungen an die Gestaltung von Eröffnungen in der HRI diskutiert.

2. KO-KONSTRUKTION UND KONTINGENZ: ERÖFFNUNG EINES FOCUSED ENCOUNTER

2.1 Situiertheit und Kontingenz im ›Interaktionssystem Mensch-Roboter‹

Entsprechend der grundlegenden Situiertheit und Kontingenz menschlicher Kommunikation ist ihr Handlungsablauf nicht im Vorfeld planbar, sondern entwickelt sich lokal, Schritt für Schritt als gemeinsame und wechselseitige Aushandlungsleistung aller Beteiligten. Diese Offenheit zieht als praktische Aufgabe für alle Beteiligten nach sich, dass eine permanente Anforderung an ihr Handeln darin besteht, interaktive Anschließbarkeit (»*what's next?*«) herzustellen. Dazu interpretieren sie vorangehende Beiträge im Hinblick auf relevante *next moves* und implementieren in der Gestaltung eines Beitrags mehr oder weniger explizite strukturelle Angebote für relevante Folgehandlungen (z.B. über projektive Strukturen unterschiedlicher Reichweite, siehe Streeck/Jordan 2009, Auer 2009, Dausendschön-Gay/Krafft 2009). So weist Levinson (2006) in der Formulierung von Grundzügen einer »*Interaction Engine*« Sequenzstrukturen als Basis von Interaktion aus, die nicht auf Regeln, sondern auf Erwartungen basieren: »Interaction is characterized by action chains and sequences (Schegloff 2006) governed not by rule but by expectations«. Daher sei insgesamt keine »*formal grammar of discourse*« möglich (Levinson 2006: 45). Während Levinson dieses zunächst auf die Beschreibung von menschlicher Kommunikation bezieht, stellt sich hinsichtlich der Systematisierung und Modellierung von Dialog- und Interaktionsverhalten für technische Systeme die Frage, wie mit der Situiertheit, der Kontingenz und dem Management von Erwartungen umgegangen werden kann.

Traditionelle Ansätze zur Modellierung von Kommunikationsverhalten technischer Systeme für die Mensch-Maschine- bzw. Mensch-Roboter-Interaktion basieren auf planbasierten Handlungslogiken, bei denen idealisierte Erwartungen über Nutzerverhalten als Basis für fixe Ablaufpläne und Zustände des Systems fungieren. Am Beispiel der Bedienung eines Kopiergerätes deckt Suchman's (1987) wegweisende Studie die grundlegende Diskrepanz zwischen solchen »*plans*« der Maschine und den menschlichen »*situated actions*« auf. Darauf reagierend werden bei der anschließenden Weiterentwicklung von technischen Modellierungsansätzen grundlegende Prinzipien wie z.B. lokales Dialogmanagement (Frohlich/Luff 1990), Inkrementalität (Skantze/Schlangen 2009) etc. berücksichtigt. In dem Maße, in dem in den Anfängen dieser Entwicklung auch explizit der Ethnomethodologie/ Konversationsanalyse nahe stehende Forscher involviert waren, hat die Beschäftigung mit Fragen des Dialogdesigns auch zur Diskussion von konversationsanalytischen Konzepten und Begrifflichkeiten angeregt. Dabei erliegen Frohlich/Luff (1990) der Versuchung, die Beobachtungen zu *Regelhaftigkeiten* von Verfahren in menschlicher Kommunikation als zu implementierende *Regeln* für technische Systeme zu verstehen. In seiner notwendigen Dekonstruktion dieser programmatischen Ideen als »*blind alley*« expliziert

Button (1990) den Unterschied zwischen computationell-deterministischen Regeln einerseits und beobachtbaren Regularitäten in menschlicher Interaktion und ihrem Status als »*practice or usage*« (Schegloff 1992), an deren Strukturen sich Teilnehmer *orientieren*. Für derartige Missverständnisse boten sich frühe Formulierungen zur »*machinery*« und »*rules of conversation*« an, die parallel auch von Schegloff (1992) u.a. in Diskussion mit Searle (1992) reformuliert wurden. Ähnlich konzeptuell ausgerichtet ist auch Schegloff's Kritik an der Rezeption von konversationsanalytischen Forschungsergebnissen beim Design von Dialogfähigkeiten in der Künstlichen Intelligenz-Forschung. Am Beispiel »Interaktionseröffnung« dekonstruiert er die Vorstellung von verbalen Eröffnungspraktiken als »Routinen«, d.h. als fixen Ablaufskripten, und expliziert – für den Gegenstand »menschliche Kommunikation« – ihren interaktiven Herstellungscharakter.

Diese konzeptuelle, auch heute nach wie vor aktuelle Diskussion rückt die zentrale, in der HRI allerdings bislang selten explizit diskutierte Frage in den Fokus, worauf genau die Modellierung zielt (vgl. Pitsch, accepted). Sollen menschliche Interaktionsfähigkeiten und -verhalten möglichst 1:1 *reproduziert* werden (z.B. Kanda/Ishiguro 2012) oder sollen möglichst *funktionale* Lösungen für die Bedienschnittstelle »Mensch-Roboter-Interaktion« geschaffen werden (z.B. Breazeal 2003)? – Wir greifen im Folgenden den Ansatz Suchman's (1987: 192) auf, demzufolge »interface design [is ...] less a project of simulating human communication than of engineering alternatives to interaction's situated properties«. Hierfür dienen uns Erkenntnisse aus der Mensch-Mensch-Interaktion (HHI) als Inspiration und Startpunkt, da sie auch zum Erfahrungsrepertoire der menschlichen Nutzer gehören. Gleichzeitig besteht die zentrale Frage darin, wie diese in der HRI funktionieren, welche Implikationen daraus erwachsen und wie sie in interdisziplinärer Kooperation und einem iterativen *research loop* verändert, erweitert oder ggf. verworfen werden sollten. Diese Herangehensweise ermöglicht insbesondere auch neue Perspektiven als Basis für die Modellierung, wie z.B. danach zu fragen, wie Nutzer das Verhalten des Roboters jeweils lokal interpretieren und wie ihre Erwartungen an die Kompetenzen des Roboters und an relevante Anschlussmöglichkeiten im Interaktionssystem lokal durch das Verhalten des Roboters selbst sukzessive beeinflusst und pro-aktiv mitgestaltet werden können (vgl. Pitsch et al. 2012, 2013, accepted).

Den konzeptuellen Ausgangspunkt bildet hierfür die Betrachtung von Mensch und Roboter als *ein Interaktionssystem* (Luhmann 1984), das gemeinsam die anstehenden Aufgaben bearbeitet und dafür Zuständigkeiten und Funktionalitäten im Interaktionssystem entsprechend der Möglichkeiten seiner Mitglieder verteilt und organisiert. Diese interaktionistische Sicht findet ihre Entsprechung im soziologischen Konzept »hybrider sozio-technischer Konstellationen« (Rammert/Schulz-Schaeffer 2002), die in Mensch-Maschine-Konstellationen die jeweiligen Beiträge zur Bewältigung einer Aufgabe als »Mit-Handeln« technischer Artefakte und als »Mit-Funktionieren« menschlicher Akteure konzeptuell fassen. Damit wird »Intelligenz« nicht als eine individuelle Leistung verstanden, sondern als eine verteilte sozio-technische Größe, für deren Definition die Zuschreibungsprozesse der Interaktionsteilnehmer und ihre gesellschaftlichen Deutungspraktiken zentral sind. Die Frage nach der *agency*/ Handlungsfähigkeit technischer Systeme wird damit zu einem empirisch beobachtbaren Phänomen (vgl. Krummheuer 2010).

2.2 Eröffnung eines *focused encounter* in HRI und HHI

Der Einstieg in ein *focused encounter* in der Mensch-Maschine-Interaktion stellt insbesondere dann einen zentralen Moment dar, wenn die Nutzer einer neuen Technologie – wie in unserem Fall: ein humanoider Roboter – zum ersten Mal mit dieser in Kontakt treten. Die menschlichen (potentiellen) Nutzer stehen vor der Aufgabe herauszufinden, welche Funktion das Gerät hat, über welche Kompetenzen das System verfügt, inwiefern es responsiv auf ihre Handlungen reagiert und was in dieser Konstellation relevante Handlungsoptionen für den Nutzer darstellen. In Studien zur HRI haben solche Einstiegssituationen jedoch kaum explizit Beachtung gefunden. Zum einen wurde ein stärkerer Fokus auf das physische Erscheinungsbild gelegt, verbunden mit der Annahme, dass insbesondere hierüber Kompetenzzuschreibungen der Nutzer erfolgen (Hegel et al. 2009) und dass vorgängige *pre-conceptions* der Nutzer so stark sind, dass sie sich nicht durch das Roboterverhalten beim Einstieg verändern (Fischer 2006). Zum anderen erproben HRI-Studien vielfach in kontrollierten Labor-Situationen spezifische Funktionalitäten eines Systems, für das diese u.U. hinsichtlich der allgemeinen Funktionsweise des Roboters in einem vorangehenden Training über die Roboterkompetenzen informiert werden (Lohse 2009). In realweltlichen Situationen wird u.U. das Eröffnungsverhalten durch einen menschlichen *Wizard* ferngesteuert (Shiomi et al. 2010). In den Fällen, in denen ein Interaktionseinstieg modelliert wird, beinhaltet dieser zumeist einen fest vorgegebener Handlungsablauf, der den aus der HHI bekannten Schritten – Gruß, Identifikation, Wie geht's, Formulierung eines Topic – folgt. Dabei stellt Lohse (2009) u.a. fest, dass solche Routinen hilfreich für den Nutzer sein können, weil sie ihm strukturelle Orientierung bieten, was insbesondere bei auftretenden Schwierigkeiten mit einem autonomen Robotersystem relevant sein kann. Demgegenüber geben erste Versuche, die Möglichkeiten für dynamischeres roboterseitiges Interaktionsverhalten zu explorieren, Hinweise darauf, dass die Art und Weise, wie Nutzer das Interaktionsverhalten des Roboters in den ersten Sekunden ihrer Begegnung erleben, einen Einfluss auf ihre weiteren Verhaltensweisen in der Interaktion mit dem Roboter hat: In einem Tutoring-Szenario (Laborstudie) hat die initiale Responsivität des Roboters Auswirkungen auf den Präsentationsstil des menschlichen Tutors (dialogisch vs. monologisch, Pitsch et al. 2012). In einem real-weltlichen Museumsszenario verbleiben Besucher, die durch ein ›*pause & restart*‹-Verfahren den Roboter als responsiv erleben, im Vergleich zu Besuchern mit non-responsivem Einstieg länger in der Interaktion und zeigen soziale Verhaltensweisen (Pitsch et al. 2009). Auch eine erste Exploration von distanzbasiert responsivem Blickverhalten des Roboters im *pre-beginning* gibt, in einer Video-Fragebogen-Studie, erste Hinweise darauf, dass die Nutzer den Roboter als kontaktinteressiert und verfügbar bewerten (Holthaus et al. 2011).² Diese Beobachtungen legen die Frage nahe: Kann man Erkenntnisse zum *pre-beginning* in Einstiegssitua-

2 Parallel hierzu erfolgt die Arbeit an den Beobachtungskompetenzen des Roboters, für die zum Erkennen von ›engagement intentions‹ die Parameter Distanz, Aufmerksamkeitsorientierung und Bewegungstrajektorien des (potentiellen) Nutzers einbezogen werden (z.B. Bohus et al. 2009).

tionen in HHI fruchtbar machen für das Design von Eröffnungssituationen in HRI? Welche Implikationen erwachsen aus einer solchen Gestaltung für die gemeinsamen Herstellungsmöglichkeiten und Ko-Konstruktion von Mensch und Roboter in der Eröffnungssituation?

In jüngeren multimodal orientierten konversationsanalytischen Untersuchungen zur Interaktionseröffnung (Mondada 2009; Mortensen/Hazel 2014; Kendon/Ferber 1973) werden besonders die wechselseitigen Koordinierungsleistungen im *pre-beginning* sichtbar, durch die Aufbau von Wechselseitigkeit und Verfügbarkeit, reziproke Identifizierung, wechselseitige Kategorisierung, Definition der Interaktionsmodalitäten und das Herstellen eines gemeinsamen Interaktionsraumes geleistet werden. Diese (zum Zeitpunkt der Studie: Mondada 2009, Kendon/Ferber 1973, sowie eine eigene HHI-Studie im Museum) dienen als Basis für das Design des Eröffnungsverhaltens des Roboter-Systems. Dieses sind insbesondere: (a) bei der Annäherung die Blickzuwendung in der Distanz sowie eine auf den potentiellen Interaktionspartner ausgerichtete Gang-/Bewegungstrajektorie, die Projektabilität erlauben hinsichtlich einer bevorstehenden Kontaktaufnahme und Kategorisierung des Gegenüber (Mondada 2009, Mortensen/Hazel 2014). (b) beim Abschluss der Eröffnungsphase und Übergang zum *first topic* die Rekonfiguration des Interaktionsraums und das Herstellen von physischen Bedingungen für die anstehende inhaltlich-thematische Arbeit.

Erprobt man solche Verfahren als roboterseitige Verhaltensweisen im *pre-beginning*, stellt sich die Frage, wie die Nutzer das Verhalten des Roboters im Hinblick auf die Interaktionskompetenzen des Systems interpretieren. Welche Möglichkeiten bieten die implementierten Handlungsstrukturen für *interactional achievement* und Ko-Konstruktion in der Mensch-Roboter-Interaktion? Welche Implikationen erwachsen daraus für die Fortsetzung der Interaktion?

3. STUDIE: ROBOTER ALS MUSEUMSFÜHRER

3.1 Design von Roboteraktivitäten und Interaktionsverhalten

Im Sommer 2010 haben wir – als Teil eines längerfristigen interdisziplinären Forschungsprojekts – eine erste explorative Studie mit dem humanoiden Roboter NAO durchgeführt, der, konfiguriert als autonomes System, in der Kunsthalle Bielefeld zufällig vorbeikommende Besucher begrüßt und ihnen Informationen zu drei Exponaten anbietet, bevor er sie wieder verabschiedet (Pitsch/Wrede 2014). Dabei lag ein besonderer Fokus auf dem Design des roboterseitigen Eröffnungsverhaltens. Um Möglichkeiten interaktiver Dynamik zu erproben, wurden Erkenntnisse aus der HHI zum *pre-beginning* aufgenommen und den traditionell-verbalsprachlichen vordefinierten Begrüßungsroutinen vorangestellt. Durch diese (u.a. auch technisch bedingte) Diskrepanz von initialer Dynamik und folgenden festen Ablaufstrukturen entstehen – so unsere Annahme – Momente, an denen das Roboterverhalten mit den anfänglich aufgebauten Erwartungen der Nutzer bricht, so dass hier im Sinne Garfinkel'scher Krisen-Experimente (Garfinkel 1967) über das Beobachten von Irritationen die

menschlichen Zuschreibungsleistungen sichtbar und ihre Verfahren zum Umgang damit beschreibbar werden.

Das Roboter-System verfügt über die folgenden Ressourcen: (1) basale Möglichkeiten zur Beobachtung von sich verändernder Position und Kopforientierung der Besucher im Raum, die über ein externes infrarotbasiertes Motion Capture System (Vicon) realisiert werden; (2) ein Set an multimodalen verbal-gestischen Roboter-Äußerungen, die vom System entsprechend eines vordefinierten Ablaufplans abgespielt werden; (3) Möglichkeiten zu *kleinschrittig-responsiver* Kopfausrichtung, die sich jeweils auf den am nächsten stehenden Besucher orientiert und dessen wechselnde Position im Raum mitverfolgt (im *pre-beginning* kontinuierlich, in der weiteren Erläuterung jeweils am Äußerungsende); (4) eine Repräsentation des Ausstellungsraumes, die diesen in 3 distanzbasierte Zonen von unterschiedlicher Interaktionsrelevanz gliedert (siehe oben; vgl. Holthaus et al. 2011).

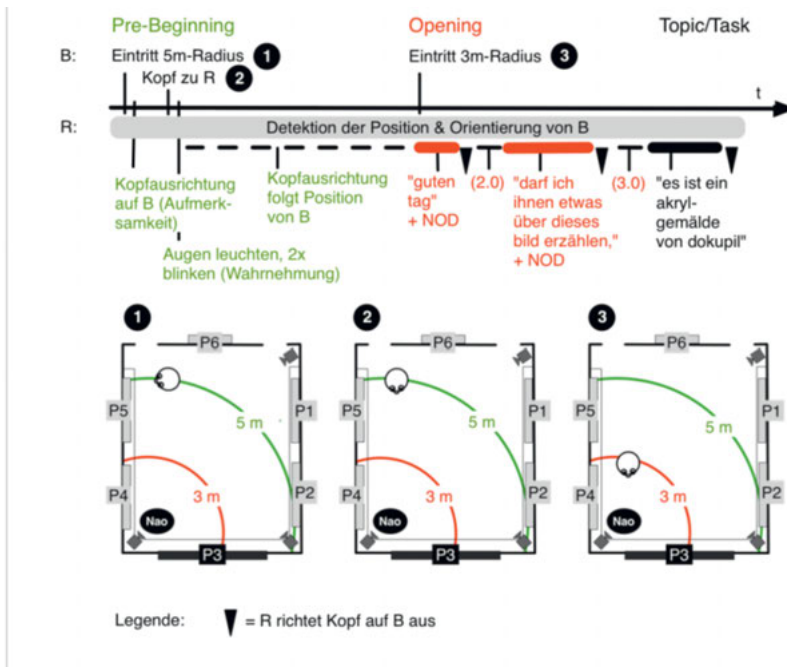
Auf dieser Basis wurden die folgenden Verhaltensweisen implementiert:

- *Pre-beginning*: Wenn ein Besucher die äußere Kreismarkierung (ca. 5m) betritt, d.h. von Zone 3 in Zone 2 eintritt, dann richtet der Roboter seinen Kopf auf die detektierte Person aus (Abb. 1, Szene 1). Hierdurch soll – so unsere Annahme – Aufmerksamkeit und *availability* signalisierbar werden. Wenn die Person zudem den Kopf auf den Roboter ausrichtet, dann blinken dessen Augen zweimal (Abb. 1, Szene 2), was die Wahrnehmung der Besucher-Orientierung kenntlich machen soll. Anschließend vollzieht der Roboter die Bewegung des Besuchers im Raum mit, indem seine Kopfausrichtung kontinuierlich dessen beweglicher Position folgt.

- *Opening*: Wenn der Besucher die innere Kreismarkierung (ca. 2m) betritt, d.h. von Zone 2 in Zone 1 eintritt (Abb. 1, Szene 3), dann begrüßt der Roboter verbal mittels »guten tag« und Kopfnicken, pausiert für ca. 2 Sekunden und fährt fort mit dem Angebot »darf ich ihnen etwas über dieses bild erzählen«, wobei er am Ende der Äußerung erneut nickt. Nach weiteren ca. 3 Sekunden setzt er mit der inhaltlichen Erläuterung (*first topic*) fort »es ist ein akrylgemälde von dokupil«. Die Beobachtungen von Heath (1992) zur Reziprozität von Nicken aufgreifend, stellt die Implementierung des Kopfnickens einen Versuch dar, die Relevanz multimodalen Interaktionsverhaltens für diese Situation zu signalisieren und zu erproben, inwieweit Nutzer solche Kopfgesten – oder weiter gefasst: Verfahren natürlicher Kommunikation – ggf. als kommunikative Ressourcen wieder aufgreifen und in der Interaktion mit dem Roboter zu verwenden suchen.

Im Anschluss an den Einstieg werden zwei Exponate vom Roboter erläutert (P3, P6) und allgemeine Informationen über die Künstlergruppe angeboten. Dabei folgen die multimodalen Äußerungen (Sprache, Gestik), ihre Abfolge und jeweiligen verbalen Pausen einem vordefinierten Ablaufplan, der in einer *Finite State Machine* modelliert und vom Nutzer nicht beeinflussbar ist. Lediglich die Kopforientierung des Roboters richtet sich weiterhin dynamisch am Ende einer Äußerung auf den ihm am nächsten stehenden Besucher aus. Für die Besucher ist diese Kombination aus fixen Ablaufstrukturen und lokaler Dynamik nicht von vornherein als solche erkennbar, sondern offenbart sich im Verlauf der Interaktion.

Abbildung 1: Interaktionseinstieg. Design des Roboter-Verhaltens



3.2 Studie und Daten

Im Sommer 2010 haben wir eine Studie mit dem beschriebenen Roboter-System (Aldebaran Nao, siehe Standbild 27.24 in Abschnitt 6.1) in der Kunsthalle Bielefeld durchgeführt (Pitsch/Wrede 2014). Während der Studie wurde der Roboter auf einem kleinen Tisch in der Ecke eines 5 x 6 m großen Ausstellungsraums positioniert. Als Nutzer des Systems fungierten die Besucher der laufenden Kunstausstellung, die kein primäres Interesse an Technologie oder Robotern haben. Sie wurden vor Betreten des spezifischen Ausstellungsraums auf die Durchführung der Studie hingewiesen und gebeten, dazu die vom Forschungsteam bereit gestellten Mützen (auf denen sich die entsprechenden Erkennungsmarker für das Infrarot-Tracking befinden) zu tragen. Sie erhielten keine weiteren Informationen über den Inhalt der Studie, bekamen im Nachhinein aber umfänglich Gelegenheit zu Fragen und Erklärungen durch das Forscherteam. Die Studie fand während einer Woche mit insgesamt 260 Gruppen statt und wurde mit 2 bzw. 3 Videokameras gefilmt. Die im Folgenden verwendeten drei Fallbeispiele wurden aus dem Gesamt-Corpus ausgewählt, weil die Interaktionsbeteiligten gut erkennbar sind, der Roboter ein relativ konsistentes Verhalten zeigt, wenig zusätzliche Ablenkungen auftreten und sie – entsprechend des explorativen Charakters der Analyse – komplementäre Perspektiven auf die Frage der Ko-Konstruktion beim »Einstieg in ein *focused encounter*« erlauben.

4. ZUGÄNGLICHKEIT ALS PRAKTISCHES PROBLEM: ALLTAGSTECHNOLOGIE, NATÜRLICHSPRACHLICHE BEDIENUNG, WECHSELSEITIGKEIT?

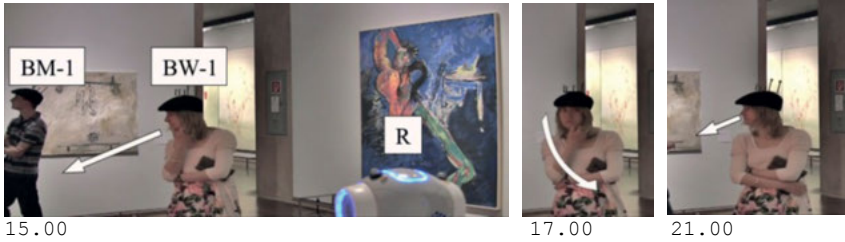
Die Untersuchung der Frage, wie Mensch und Roboter beim Einstieg in ein *focused encounter* Zugänglichkeit herstellen, beginnen wir mit einem Fallbeispiel (VP 081), bei dem sich dies als ein praktisches Problem erweist: Die Besucherin zögert zwischen der Verwendung von Mitteln natürlichsprachlicher Kommunikation und der Bedienung einer herkömmlichen Alltagstechnologie mittels einer enaktierten Fernbedienung. An diesem Beispiel wird ferner bezüglich der Ko-Konstruktion in der Mensch-Roboter-Interaktion sichtbar, dass bei einer rein verbalsprachlichen Eröffnung mit vorgefertigten Ablaufstrukturen die Teilnehmer zwar die anstehenden konditionellen Relevanzen bedienen und so die Strukturen der Eröffnungsroutine abarbeiten. Dabei werden allerdings die über diese Strukturen normalerweise erfüllten Funktionen – wie z.B. das Herstellen von Reziprozität – nicht notwendigerweise auch mit bereitgestellt, so dass im weiteren Verlauf der Interaktion eine Kopfabwendung des Roboters im Zweifelsfall als Abwahl der aktuellen Beteiligungsstruktur interpretiert wird und die Besucherin anschließend als aktive Mitwikerin aus dem Geschehen aussteigt.

4.1 Fallbeispiel 1: Analyse

Wir steigen in die Analyse ein, nachdem die zwei Besucher BM-1 (männlich) und BW-1 (weiblich) den Ausstellungsraum betreten haben und sich mit Blick auf die (vom Roboter aus gesehen) an der linken Wand befindlichen Bilder dem Roboter bis auf ca. 2 m genähert haben.

(1) Herstellen von Aufmerksamkeit und Ausbleiben des Signalisierens von Wahrnehmungswahrnehmung: Als sich die Besucherin BW-1 einen weiteren Schritt annähert, richtet der Roboter seinen Kopf mit einer Linksdrehung auf sie aus (#15.00) und begrüßt sie mit »guten tag« + Kopfnicken (01), was als *summons* die Aufmerksamkeit von BW-1 auf den Roboter lenkt (#17.00). Als keine weiteren Aktivitäten des Roboters folgen, wendet sich BW nach 2 Sekunden wieder zum Exponat zurück (#21.00).

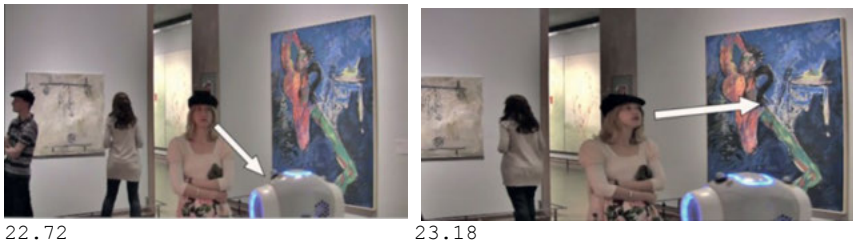
```
01 R-head: |guten tag | (2.0) | (2.0) |
    R-act: |nod      |
    BW-act: |@P4      |@R    |≈≈≈|@P4
           |15.00    |17.00  |21.00
```



D.h. der Roboter stellt durch die Begrüßung eine Situation veränderter Aufmerksamkeitsorientierung bei der Besucherin her, durch die ein struktureller *Slot* und ein Zeitfenster entstehen, in dem vonseiten der Besucherin eine Reaktion des Roboters auf die veränderte Partizipationskonstellation bzw. das Signalisieren von Wahrnehmungswahrnehmung relevant wären.

(2) Bearbeiten von körperlichen Handlungsangeboten: Ca. 2 Sekunden nachdem sich BW-1 wieder zu den Exponaten orientiert hat, setzt der Roboter mit dem Angebot einer Bilderläuterung fort – »darf ich ihnen etwas über dieses bild erzählen,« (02) – und gibt sich dadurch in der Rolle eines Museums-Guide zu erkennen. Er nickt dabei mit dem Kopf und vollführt mit dem rechten Arm eine Zeigegeste zum Bild P3 (vgl. Pitsch/Wrede 2014). BW-1 reagiert hierauf, indem sie – bereits während der laufenden Äußerung Naos – dessen primär *körperlichen* Handlungs- bzw. Koordinierungsangeboten folgt und deren interaktive Implikationen sukzessive abarbeitet: Im Anschluss an »darf ich ihnen« wendet sie sich erneut dem Roboter zu (Herstellen von Aufmerksamkeit), tritt einen Schritt näher an ihn heran und positioniert sich dabei in dessen vermeintlichem Blickfeld (#22.72, Konstitution von Interaktionsraum) und folgt anschließend dem Orientierungshinweis auf das Bild per Kopf- und Körperdrehung (#23.18, Herstellen von Referenz).

```
02 R:      |darf ich ihnen etwas |über |dieses |bild |erzählen, |
R-act:    |                                     |nod      |
R-ges:    |onset                               |deix-to-P3 |retract |
BW-gaz:   |P4                               |@R        |~~~~~  |@P3
          |22.72                           |          |23.18
```



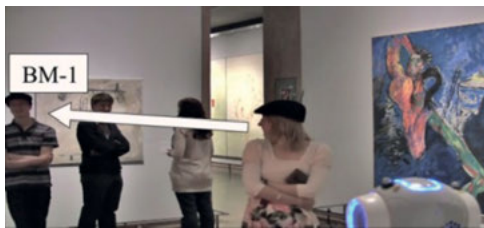
(3) Natürlichsprachliche Kommunikation vs. herkömmliche Alltagstechnologie: Beim anschließenden Beantworten der *verbal-expliziten* Ebene der Frage aber zögert BW-1. Sie dreht sich zu ihrem Begleiter BM-1 um und fragt: »soll ich jetzt mit JA« (03, #25.09), unterbricht ihre entstehende Äußerung, wendet sich wieder zum Roboter zurück und antwortet lachend »ja erzähl mal« (04, #27.24), was parallel

auch von BM-1 mit »ja natürlich;« (04) bestätigt wird. Unmittelbar anschließend richtet sie ihren ausgestreckten Arm auf Nao und imitiert gestisch mit dem Drücken eines Knopfes und lautlich mit der Verbalisierung eines »klick«-Geräusches das Bedienen einer Fernbedienung (04, #29.04 a+b). In dem Maße, in dem der Roboter unmittelbar mit der inhaltlichen Erläuterung »es ist ein akrylgemälde von dokupil« (05) fortsetzt, werden die von BW-1 unternommenen Aktivitäten als relevante, für den Roboter anschlussfähige Handlungen im sich formierenden »Interaktionssystem Mensch-Roboter« bestätigt. Die Besucherin quittiert dies mit sichtlicher Freude.

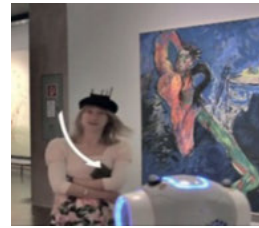
03 P: (1.5) |
 BW: |soll ich jetzt mit JA-|(-)
 BW-gaz: |@R
 |25.09

04 BM: |ja na|türlich;
 BW: |ja | |erzähl mal |(klick)
 BW-act: |fernbedienung |
 R-head: |@BW |-->
 |27.24 |28.12 |29.04 |30.16

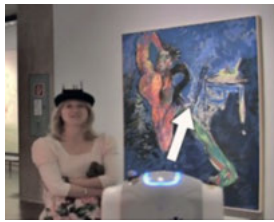
05 R: |es ist ein akrylgemälde von dokupil;
 BW-gaz: |@P3



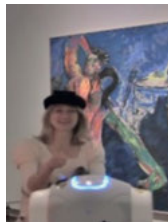
25.09



27.24



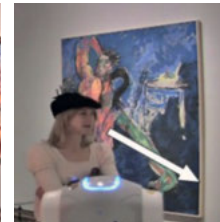
28.12



29.04a



29.04b



30.16

Mit dieser Abfolge aus Zögern und der Ambivalenz zwischen einer Antwort mittels Verfahren natürlichsprachlicher Kommunikation, wie sie vom Roboter selbst angeboten werden, und dem Rückgriff auf tradierte Verfahren des Bedienens herkömmlicher Alltagstechnologien gibt die Besucherin ihre Unsicherheit darüber zu erkennen, wie sie mit dem Roboter umgehen soll, worauf er potentiell reagieren könnte und was für sie relevante Handlungsmöglichkeiten darstellen können. Das Herstellen von Anschlussfähigkeit stellt also ein manifestes praktisches Problem dar.

(4) Erproben des Modus »natürlichsprachliche Kommunikation«: Mit dem Übergang zum *topic talk* definiert der Roboter die Eröffnung als beendet. Im nächsten

Interaktionsschritt folgt die Frage des Roboters »kennen sie dokupil,« (07), was für BW-1 (und den analysierenden Forscher) Gelegenheit bietet, ihre Zuschreibungsleistungen zu überprüfen. In der Gestaltung ihrer Antwort geht BW-1 auf die vom Roboter angebotenen Verfahren natürlichsprachlicher Kommunikation (anstelle der vorherigen alltagstechnologischen Variante) ein: [»NEIN« + unspezifisches Kopfkreisen] (08). Als darauf keine unmittelbare Reaktion des Roboters folgt, reformuliert sie ihre Antwort als ein ostentatives Kopfschütteln und nimmt damit die vom Roboter angebotene Kopfgestik als kommunikative Ressource auf (08). D.h. einerseits lässt sie sich auf den Modus »natürlichsprachliche Bedienung« ein, was eine zentrale Leistung der Eröffnung darstellt. Andererseits sind die Details der im »Interaktionssystem Mensch-Roboter« relevanten Handlungsoptionen für sie noch nicht transparent und müssen während der inhaltlichen Arbeit weiter erprobt werden.

06 (1.5)

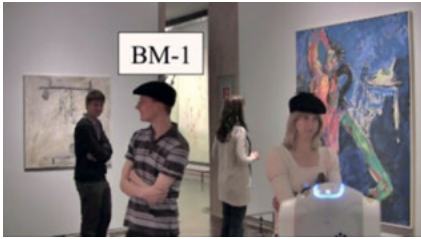
```
07 R:          kennen sie |dokupil,|
   R-act:      |nod nod |

08 BW:        |NEIN      | (1.3) | (1.0) |
   BW-act:    |H-kreist  |      |H-shake|
```

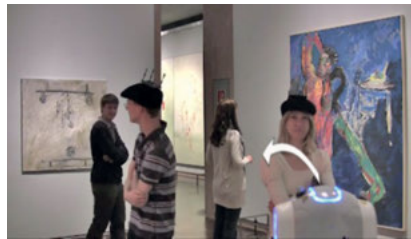
(5) Kopfabwendung interpretiert als Ausbleiben von Reziprozität und Abwahl des Interaktionspartners: Anschließend gelingt es der Dyade allerdings nicht, an der Interaktionsoberfläche weiterhin Wechselseitigkeit von Handlungen herzustellen. Insbesondere behandelt BW-1 eine sich verändernde Kopfausrichtung des Roboters als Abwahl ihrer Person aus der Rolle des primären Interaktionspartners, so dass sie anschließend als aktive Mitwirklerin aussteigt und sich eine veränderte Teilnehmerkonstellation ergibt. – Konkret tritt während BW-1's Antwortversuch (08) auch BM-1 näher an den Roboter heran (#39.68), was Nao dazu veranlasst, im Anschluss an BW-1's zweiten Antwortversuch seinen Kopf auf BM-1 auszurichten (#40.40). BW-1 folgt – mit sichtlichem Erstaunen – diesem veränderten (vermeintlichen) Aufmerksamkeitsdisplay des Roboters und dreht sich zu BM-1 um (#42.00). Als sie wieder zu Nao zurückblickt, setzt dieser mit einer Erläuterung zu Dokupil fort (09, #43.00). D.h. auf sequenzstruktureller Ebene behandelt der Roboter ihre Aktivitäten nicht als relevante Anschlusshandlungen, sondern – im Gegenteil – wendet sich BM-1 zu. Dieses behandelt BW-1 im Folgenden durch ihren Ausstieg als eine vom Roboter herbeigeführte Veränderung in der Partizipationsstruktur.

```
08 R-head:      |@BM |
   BW:         |NEIN | (1.3) | | (0.5) |
   BW-act:     |H-kreist | |H-shake|
   BW-gaz:     |@R      |@BM |
               |39.68 |40.40 |42.00

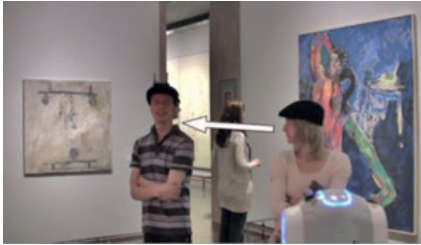
09 R:          |dokupil ist ein neoexpressionistischer maler
   BW-gaz:     |@R
               |43.00
```



39.68



40.40



42.00



43.00

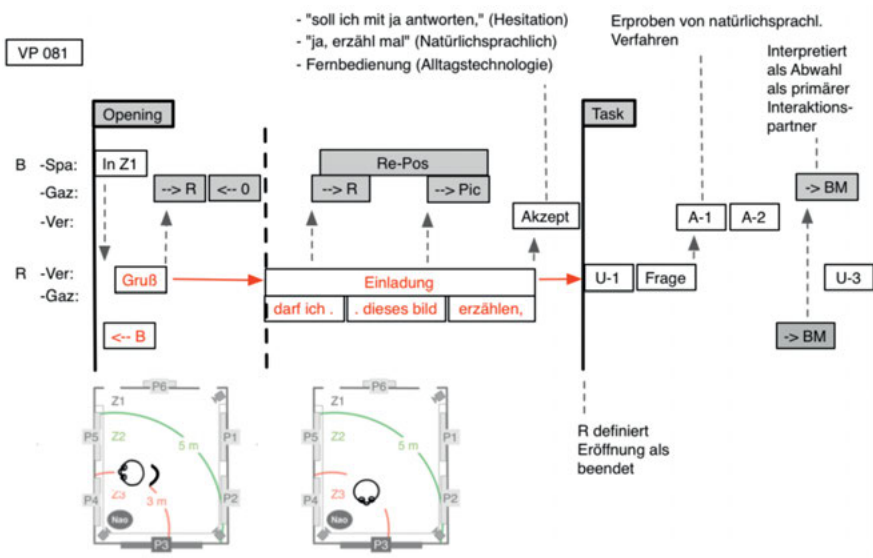
D.h. nach der beginnenden Herausbildung eines Interaktionssystems ›BW-1 – Roboter‹ und dem Erproben der möglichen anschlussfähigen Handlungsmodalitäten findet hier – zumindest für BW-1 – keine Stabilisierung der Teilnehmerkonstellation und der interaktiven Anschlussmöglichkeiten statt.

4.2 Fazit: Ko-Konstruktion – Abarbeiten von Strukturen vs. Erledigen kommunikativer Funktionalitäten

Betrachtet man diesen Einstieg und beginnenden Interaktionsverlauf vor der Folie des Konzepts von Ko-Konstruktion, dann wird die Diskrepanz zwischen dem Abarbeiten von verbalsprachlichen Strukturen und dem Erledigen tatsächlicher kommunikativer Aufgaben und Funktionen manifest (Siehe Abb. 2). Es wird vom Roboter eine feste Abfolge von verbalsprachlichen Eröffnungsroutinen (Gruß, Frage) vorgegeben, deren starke strukturelle Anforderungen jeweils von der Besucherin abgearbeitet werden. Dabei muss sie zunächst herausfinden, was im Interaktionssystem geeignete relevante Anschlusshandlungen darstellt. Entsprechend des vorprogrammierten Handlungsablaufs arbeitet der Roboter sukzessive auf die Beendigung der Aufgabe ›Interaktionseröffnung‹ hin und deklariert sie mit dem Einstieg in den *topic talk* als beendet. Das heißt, die gemeinsame Herstellungsleistung in diesem Mensch-Roboter-Interaktionssystem besteht im roboterseitigen Angebot von Handlungsinitiativen, deren konditionelle Relevanz jeweils von der Benutzerin bedient wird. Die Beendigung der Aufgabe ist somit nicht das situative Ergebnis einer tatsächlichen wechselseitigen Aushandlung, sondern stellt eine situationsunabhängige strukturelle Setzung durch die Handlungslogik des Roboters dar. Für die Besucherin sind damit noch nicht notwendigerweise auch alle – für gewöhnlich in einer Eröffnung hergestellten – Aufgaben bearbeitet und die damit verbundenen interaktiven Funktionen etabliert. Dieses zeigt sich in Bezug auf die Frage des Herstellens von Wechselseitigkeit und adäquaten Anschlusshandlungen insbesondere im Umgang der Besucherin

mit der Blickabwendung des Roboters. Diesen interpretiert sie als eine vom Roboter herbeigeführte Veränderung in der Partizipantenkonstellation, die den Beginn ihres Ausstiegs aus der Rolle der aktiven Mitwirkenden hin zu einer Beobachterin der sich dann neu formierenden Dyade markiert.

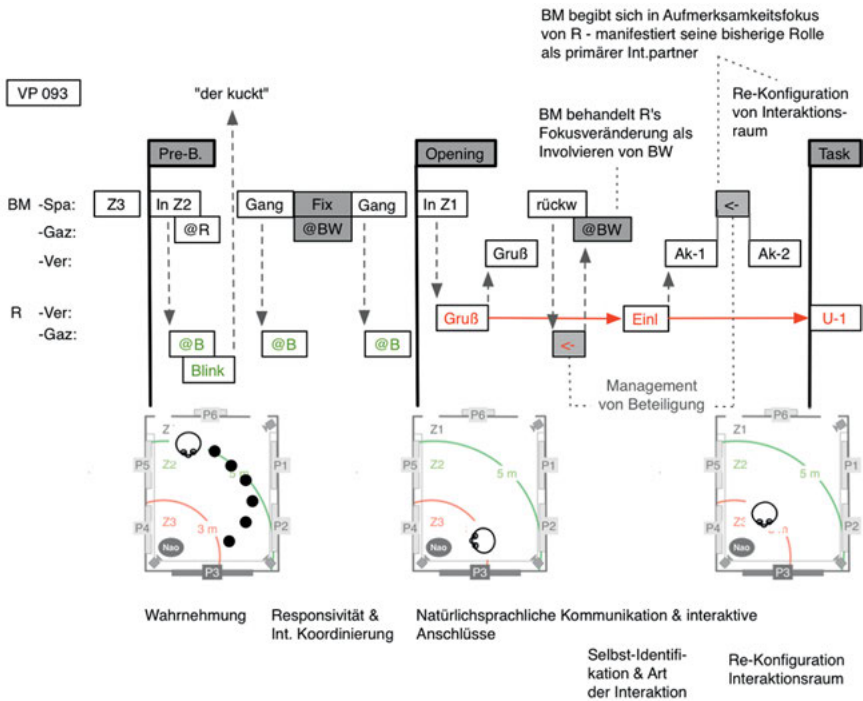
Abbildung 2: Handlungs- und Sequenzstruktur VP 081.



5. INTERAKTIVE KOORDINIERUNG IM *PRE-BEGINNING* UND HERSTELLEN VON ANSCHLUSSMÖGLICHKEITEN

In einem zweiten Fall (VP 093) beginnt der Einstieg in das *focused encounter* mit einem *pre-beginning*, in dem der Besucher den Roboter frühzeitig – d.h. bereits vor dem Bearbeiten der verbalen Eröffnungsroutinen – als ein responsives Gegenüber erfährt (siehe Abb. 3). Interessanterweise und im direkten Vergleich zu Fallbeispiel 1 wird hier die Aktivität des roboterseitigen Kopfabwendens (die einen Interaktionsschritt früher als in Fallbeispiel 1 erfolgt) nicht als Abwahl des bisherigen Interaktionspartners interpretiert, sondern als Angebot zur Rekonfiguration von Interaktionsraum genutzt.

Abbildung 3: Handlungs- und Sequenzstruktur VP 093.



5.1 Fallbeispiel 2: Analyse

Für die Analyse steigen wir in die Interaktion ein, wenn die zwei Besucher BW-2 (weiblich) und BM-2 (männlich) den Ausstellungsraum betreten.

(1) Signalisieren von wechselseitiger Wahrnehmung: Wenn die beiden Besucher BM-2 und BW-2 den Raum betreten, detektiert das technische System auf einer Distanz von ca. 5 Metern ihre Anwesenheit und erkennt ihren Eintritt in Zone 2 (siehe oben 3.1), so dass der Roboter seinen Kopf auf sie ausrichtet (#05.20, #06.84). Nahezu gleichzeitig wendet sich BM-2 zum Roboter, während dessen Augen kurz aufblinken. Dieses wird vom Besucher als roboterseitiges Wahrnehmungsdisplay behandelt: »oh der ↑KUCKT;« (01) bzw. an seine Begleiterin adressiert »kuck mal; <name> (.) der KUCKT;« (02).

```
01 R-head: |0      |@BM/BW |0
    R-act:      |eyes blink
    BM:         |oh-   |oh der ↑KUCKT;
                |05:32 |06:84

02 BM:         (2.0) kuck mal; <name> (.) der KUCKT;
```

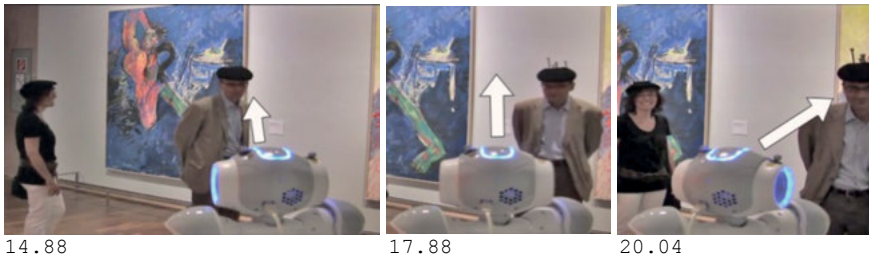


05.32

06.84

(2) Responsivität des Roboters – Kopfausrichtung folgt der Bewegung des Besuchers: BM-2 richtet nun seine Aktivitäten auf den Roboter aus und geht in den Raum hinein, wobei die Kopforientierung des Roboters seiner Bewegung sukzessive folgt (#14.88, #17.88). BM-2 quittiert dieses durch kurzes Stehenbleiben, Blickkontakt mit BW-2 und Lachen. Dann tritt BM in einer bogenähnlichen Bewegungstrajektorie näher an den Roboter heran, was dieser weiterhin per Kopfausrichtung mitverfolgt (#20.04).

03 Bm-ver: | (3.5) | OH | (2.0) | (lacht) |
| 14.88 | | 17.88 | 20.04



14.88

17.88

20.04

Insgesamt gibt sich der Roboter also für die Besucher im *pre-beginning* als ein auf visueller Wahrnehmung basierendes responsives System zu erkennen, dessen Beobachtungsmöglichkeiten BM-1 durch die Wahl seiner Gang-Trajektorie erprobt.

(3) Bedienen konditioneller Relevanzen und Modus »Natürlichsprachliche Kommunikation«: Als sich der Besucher dem Roboter auf ca. 1 Meter nähert und damit – im Modell des Roboters – die »innere Zone« betritt, löst dieses die vorprogrammierte Abfolge von Eröffnungsverhalten aus: Die verbale Begrüßung »guten tag;« + Kopfnicken (04) des Roboters wird von BM-2 mit einem ostentativen »TA::G;« + Kopfnicken (04) beantwortet. Das anschließende als Frage formulierte Angebot »darf ich ihnen etwas über dieses bild erzählen;« + Kopfnicken (05), das gleichzeitig eine Selbstidentifikation als Museumsführer und den Aktivitätstypus

›Exponat-Erläuterung‹ anbietet, wird von BM-2 mit »na klar;« (06) bedient und ca. 1 Sekunde später durch »bitte gern;« (06) ergänzt. Anschließend setzt der Roboter mit der angekündigten Erläuterung fort (07), sodass auf sequenzstruktureller Ebene BM-2's vorangehende Antworten funktional erscheinen.

```

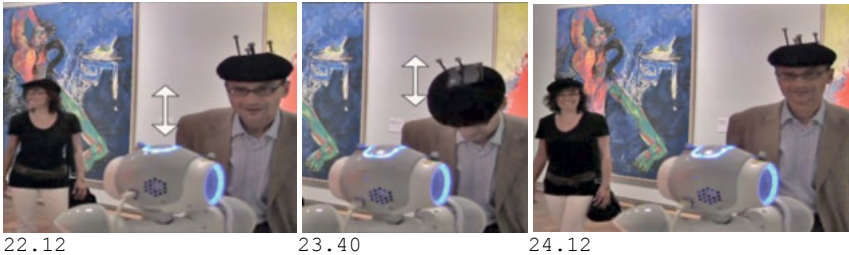
04 R:      |guten tag; |
R-head:   |nod
BM:              |TA: |:G; |
BM-act:      |nod |
            |22.12      |23.40 |24.12

05 R:      |darf ich ihnen etwas |über |dieses bild |erzählen,
R-act:     |
R-head:      |nod |

06 BM:      na klar; (1.0) bitte gern;

07 R:      es ist ein akrylgemälde von dokupil

```



Im Vergleich zum ersten Fallbeispiel entsteht hier kein Zögern des Besuchers bezüglich der Frage, wie man den Roboter bedienen könnte. Vielmehr greift er direkt die von Nao angebotenen kommunikativen Möglichkeiten auf. Hier scheinen beim Übergang von Eröffnung zum Topic Talk relevante Modalitäten, Handlungsweisen und Anschlussmöglichkeiten im Interaktionssystem ›Mensch-Roboter‹ für den Besucher bereits hergestellt zu sein.

(4) Kopfabwendung interpretiert als Angebot zur Rekonfiguration von Interaktionsraum: Im Anschluss an die Erwiderung des Grußes »TA: :G;« durch den Besucher wendet der Roboter seinen Kopf vom Besucher ab. Dieses kommt systemintern dadurch zustande, dass BM-2 nach der Erwiderung des Grußes einige Schritte zurücktritt (#24.52, #25.36) und somit die Besucherin BW-2 vom Roboter als am nächsten stehende Person identifiziert wird, so dass Nao seinen Kopf in ihre Richtung in etwa mittig zwischen beide Besucher wendet (#26.00). Diese Situation ähnelt der Situation in Fallbeispiel 1 (»kennen sie dokupil«, siehe oben 4.1 – (5)), erfolgt hier jedoch zwei *moves* früher im Interaktionsverlauf. Während dieses Roboterverhalten in Fallbeispiel 1 als mangelnde Responsivität und roboterseitige Wahl eines anderen Besuchers als primären Interaktionspartner behandelt wird, interpretiert BM-2 es hier als ein Angebot zur Rekonstitution von Interaktionsraum. – Dieses stellt sich konkret wie folgt dar: Die Frage »darf ich ihnen etwas über dieses bild erzählen,« (05) beantwortet BM-2 zunächst spontan mit »na klar« (06, #30.36), dann tritt er zur Seite in das vermeintliche Blickfeld des Roboters (#31.56)

und reformuliert seine Antwort (»bitte gern«, 06, #32.12). In diesem Sinne erscheint die emergierende Dopplung seiner Antwort responsiv auf das Roboterverhalten und die Repositionierung als Manifestation zur Aufrechterhaltung seines bisherigen Beteiligungsstatus als primärer Interaktionspartner.

```

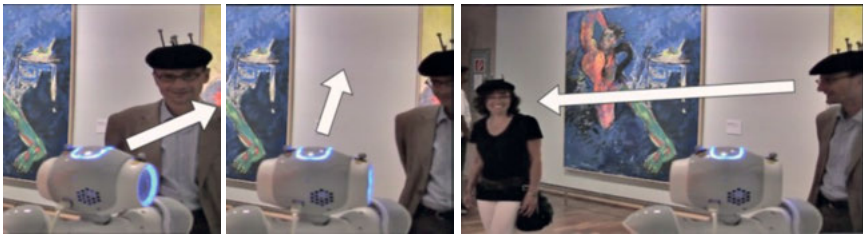
04 R:      |guten tag; |
   R-head: |nod                               |<--
   BM:      |TA: |:G; |
   BM-act:  |nod |rückwärts |
   BM-gaz:  |@R                                     |@BW
                                     |24.52   |25.36 |26.00

05 R:      |darf ich ihnen etwas |über |dieses bild |erzählen,
   R-act:  |                                     |point   |
   R-head: |nod                               |

06 BM:     |na klar; |(0.5) |(0.5) |bitte gern;|
           |30.36      |31.36 |32.12

07 R:      es ist ein akrylgemälde von dokupil

```



24.52

25.36

26.00



30.36

31.36

32.12

Das heißt, dass während im ersten Fallbeispiel die Konsequenz aus der roboterseitigen Blickabwendung das Ausscheiden von BW-1 aus der Rolle des primären Interaktionspartners war, sich BM-2 hier aktiv in eine neue Position begibt, um die bisherige Beteiligungsstruktur aufrecht zu erhalten. Damit ist gleichzeitig eine Re-Konfiguration von Interaktionsraum verbunden. Die damit eingenommene Position und entstehende räumliche Konstellation erweist sich als eine besonders geeignete, wenn der Roboter mit der inhaltlichen Erläuterung beginnt: es ist eine klassische Dreiecks-Konstellation zwischen Sprecher, Rezipient und Objekt entstanden. Damit ist – ausgehend von der Kopfwendung des Roboters – passend zum Einstieg in das *first topic* auch die – aus menschlicher Kommunikation bekannte – aufgabenbezogene Rekonfiguration des Interaktionsraumes erfolgt (siehe oben 2.2). Inwiefern die

ser unterschiedliche Umgang der beiden Besucher mit der Kopfwendung des Roboters auch mit ihren vorgängigen Erfahrungen zu Responsivität und Anschlussmöglichkeiten im System und der in der Interaktion etablierten Form von Wahrnehmung und Responsivität systematisch zusammen hängt, wäre auf einer größeren Datenbasis weiter zu untersuchen.

5.2 Fazit: Ko-Konstruktion – Herstellen von Wahrnehmung und Reziprozität im *pre-beginning*

Während das erste Fallbeispiel analog zu Suchman (1987) die Differenz zwischen »plans« und »situated actions« vorführt, erlaubt das zweite Fallbeispiel einen differenzierteren Blick auf die Möglichkeiten von Ko-Konstruktion im Interaktionssystem »Mensch-Roboter«. Mit dem Angebot von roboterseitiger Reaktivität der Kopforientierung (interpretiert als Blickverhalten) bezüglich des nutzerseitigen Raumverhaltens im *pre-beginning* entsteht eine interaktive Dynamik, in der der Besucher den Roboter zu Beginn der Kontaktaufnahme als ein responsives System erfährt. Die Handlungen von Mensch und Roboter sind aufeinander bezogen, die Gang-Trajektorie des Besuchers und seine Nutzung des Raumes emergieren im interaktiven Wechselspiel mit der Kopfausrichtung des Roboters (vgl. ähnlich zu Blick und manipulativen Handtrajektorien, Pitsch et al. 2013). Gleichzeitig bleibt im weiteren Verlauf der Interaktion die bereits im ersten Fallbeispiel herausgearbeitete Verfahrensweise des Abarbeitens der vom Roboter vorgegebenen Strukturen ebenso relevant. Auch hier definiert der Roboter durch den vorprogrammierten Übergang zum *topic talk* das Ende der Eröffnung. Allerdings erscheinen in diesem Fall für den Besucher die Handlungsmöglichkeiten im Interaktionssystem transparenter, worauf sein weiteres Interaktionsverhalten hindeutet. Auf visueller Ebene wird die Blickabwendung des Roboters nicht als Abwahl der aktuellen Partizipanten-Konstellation interpretiert, sondern als Angebot zur dynamischen Rekonfiguration von Interaktionsraum. Auf verbaler Ebene bedient der Besucher im gesamten weiteren Verlauf der Exponat-Erklärungen die strukturellen *slots*, die die Erläuterungen des Roboters eröffnen. Das heißt, bezüglich der Frage, welche Handlungsmöglichkeiten das Roboterverhalten für den Nutzer im Interaktionssystem eröffnet, lässt sich eine differenzierte Sicht des Benutzers erkennen, der die beiden gegensätzlichen, in der Eröffnung hergestellten Handlungslogiken des Systems bedient: einerseits verbales *slot-filling* und andererseits visuelle Responsivität. Letzteres ermöglicht es im Interaktionssystem über körperliche Verhaltensweisen, grundlegende Aufgaben der Interaktionsorganisation zu bearbeiten, wie z.B. die Konstitution von Interaktionsraum und das Management der Beteiligungsstruktur. Damit wird die Basis für grundlegende Funktionen des Organisierens von Interaktion gelegt und es werden Bedingungen für Stabilität im Interaktionssystem hergestellt.

6. IMPLIKATIONEN VON WAHRGENOMMENER RESPONSIVITÄT: ERWARTUNGEN AN DIE FÄHIGKEIT ZU *SECOND TURNS*

Im dritten Fallbeispiel (VP 222) machen die Besucher – ähnlich, wenngleich weniger ausgebaut – die Erfahrung eines responsiven Robotersystems im *pre-beginning*. Allerdings führt dieses anders als in der differenzierten Herangehensweise des Besuchers aus Fallbeispiel 2 (verbales *slot-filling* vs. visuelles dynamisches Management von Interaktion) hier zu der besucherseitigen Erwartung, dass der Roboter über die Fähigkeit zum Produzieren von verbalen *second turns* verfügt. Die Besucherin folgt nicht nur den verbalsprachlichen konditionellen Relevanzen des Roboters, sondern sie ergreift selbst die Initiative und leitet neue Sequenzen ein. Dies stellt erhöhte Anforderungen an die Koordinierung im Interaktionssystem, die der Roboter – in seiner aktuellen Version – noch nicht einlösen kann. Dies verweist auf die Frage, inwieweit die initial in der Eröffnung etablierten Kompetenzzuschreibungen funktional sind und in Relation zu den tatsächlichen Handlungsmöglichkeiten im Interaktionssystem stehen.

6.1 Fallbeispiel 3: Analyse

Wir steigen in die Interaktion ein, nachdem zwei Besucher BM-3 und BW-3 den Raum betreten haben und BW-3 bereits einmal so nah an den Roboter heran getreten ist, dass sie dessen Kopfwendung ausgelöst hat. Allerdings bleibt BW-3 auf die Exponate fixiert und entfernt sich wieder, sodass der Roboter den Kopf in die Ausgangsstellung zurückbewegt ohne die verbalsprachliche Eröffnungsroutine aufzuzuführen (siehe Abb. 4).

(1) Verfügbarkeit und Wechselseitigkeit von Wahrnehmung: Einige Sekunden nach dem oben wiedergegebenen ersten Anlauf nähert sich BW-3 erneut mit Blick auf die Exponate dem Roboter (#26.00), der sie wiederum als potentiellen Interaktionspartner klassifiziert und mit Kopfwendung reagiert (#26.48). Dieses erfolgt schwungvoll und dadurch mit deutlich wahrnehmbarem mechanischen Geräusch, was die Aufmerksamkeit von BW weckt und damit wie ein *summons* fungiert: Sie wendet sich zum Roboter (#27.24), der mit den Augen blinkt. BM-3 signalisiert mittels des *change-of-state token* »o:::h« Überraschung (02), lacht und interpretiert das Verhalten Nao's als roboterseitiges Display von Wahrnehmung: »ich glaub der hat mich erkannt;« (02), so dass hier der Eindruck von roboterseitiger Fähigkeit zur Wahrnehmungswahrnehmung entsteht.

01 P: |26.00 |26.48 |27.24

02 BW: o:::h, ((lacht)) ich glaub er hat mich erkannt;

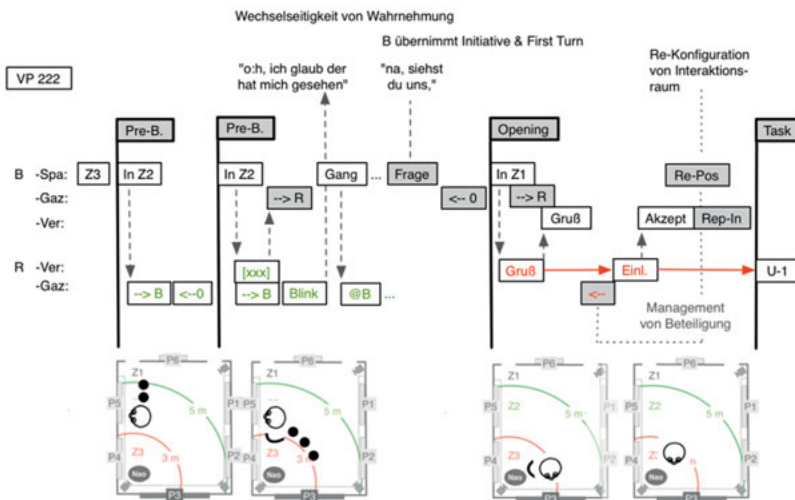


26.00

26.48

27.24

Abbildung 4: Handlungs- und Sequenzstruktur VP 222.

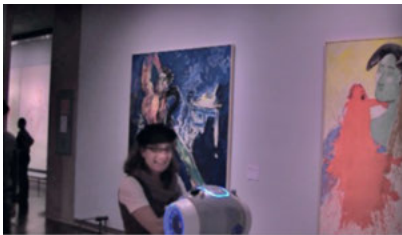


(2) **Ambivalenz interaktiver Koordinierung – Visuelle Responsivität und Ausbleiben des verbalen *second turn*:** Dann schreitet sie mit Blick zum Roboter langsam auf diesen zu, wobei sich dessen Kopfausrichtung bei jedem Schritt sukzessive mitbewegt. Anders als in Fallbeispiel VP093 ergreift nun die Besucherin die Initiative und adressiert die Frage »na:, (.) siehst du uns,« (04, #34.52, #36.36) an den Roboter, deren Design die Ambivalenz der Situation aufzeigt. Zum einen stellt es einen *first pair part* dar, der den Einstieg in einen verbalen Austausch anbietet. Zum anderen verwendet BW-3 ein Format, mit dem regelmäßig Tiere und Kleinkinder angesprochen werden und von denen man zwar keine direkte Beantwortung der Frage erwartet, wohl aber etwaige sequenziell passende Anschlüsse wohlwollend kommentiert und als Beleg von Intelligenz feiert (vgl. Bergmann 1988). Als nach ca. 1 Se-

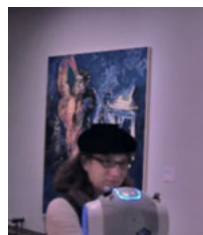
kunde keine weitere Anschlusshandlung des Roboters erfolgt ist, wendet sich BM-3 wieder ab und orientiert sich zum nächsten Exponat (#37.52). Damit behandelt sie den durch die Frage eröffneten strukturellen *slot* als nicht erfüllt und die angelegte Möglichkeit mangelnder Interaktionskompetenz des Roboters als relevant.

```
03 P:          | (3.4)
    BW-gaz:    |@R

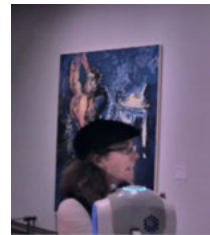
04 BW:        |na:, (.) siehst du uns, (1.0) |
    BW-gaz:    |@R                          |@P-3
                |34.52                      |36.36 |37.52
```



34.52



36.36



37.52

(3) Übernahme der Initiative durch den Roboter und Einstieg in die Routine:

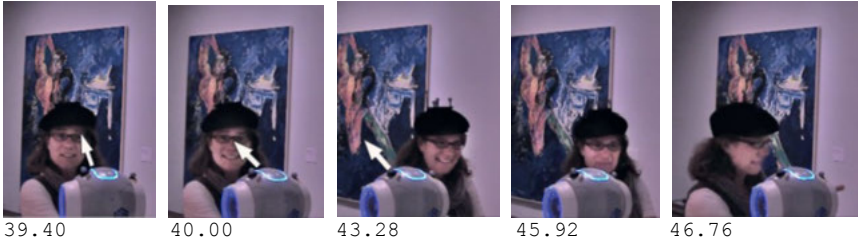
Unmittelbar nach BW-3's Umorientierung erfolgt – entsprechend des Ablaufskripts – die verbale Begrüßung des Roboters: »guten ta:g« (inkl. Kopfnicken), deren konditionelle Relevanz sie – analog zum Besucher aus Fall 2 – mit dem entsprechenden Gegengruß (»guten tag« (05)) beantwortet. Mit ihrem anschließenden lautlosen ausgedehnten Lachen (05) hält sie weiterhin die Ambivalenz aus ›Bedienen sequenzieller Strukturen‹ und ihrem Kommentieren der Situation aufrecht. Nach kurzer Pause setzt der Roboter mit dem Angebot einer gemeinsamen Tätigkeit (»darf ich ihnen etwas über dieses bild erzählen,«, 06) fort, was BW-3 mit »BITte;« positiv beantwortet. D.h. insgesamt *entsteht* hier im Interaktionssystem vorübergehend ein Modus, in dem dem Roboter die Initiative zu *first turns* überlassen wird und somit aneinander anschließende Handlungen und – an der Interaktionsoberfläche – funktionierende sequenzielle Strukturen entstehen.

```
05 R:          |guten ta:g.| (1.0) |
    R-act:      |nod          |
    R-head:     |@BW                          |<--
    BW:         |guten ↑TAG; |((lacht))
                |39.40      |40.00

06 R:          |darf ich ihnen etwas |über |dieses
    R-ges:      |pt@P1-onset
    R-head:     |nod
                |43.28

07 R:          |bild          |erzählen|
    R-ges:      |pt@P1-peak          |
    R-head:     |...                  |
                |45.92
```

08 BW: ↑BITte; (.) über DAS, |
| 46.76



(4) Kopfabwendung interpretiert als Angebot zur Rekonfiguration von Interaktionsraum im Moment des Übergangs zur inhaltlich-thematischen Arbeit: Ähnlich zu den ersten beiden Fallbeispielen gibt es auch in dieser Interaktionskonstellation einen Moment, an dem der Roboter seinen Kopf vom primären Interaktionspartner in Richtung eines anderen Besuchers abwendet. Hier erfolgt dieses an nahezu der gleichen sequenzstrukturellen Stelle wie in Fallbeispiel 2, d.h. direkt nach der Begrüßung (05, #40.00. #43.28). Analog zu Fallbeispiel 2 bearbeitet BW-3 diese Kopfabwendung zeitversetzt, und zwar erst nachdem sie die vom Roboter vorgeschlagene Einladung angenommen hat (#46.76). D.h. BW-3 behandelt die Kopfabwendung erst als ein interaktionsrelevantes Angebot zur Rekonfiguration des Interaktionsraums, nachdem die nächste gemeinsam zu bearbeitende inhaltliche Aufgabe formuliert worden ist (»darf ich ihnen etwas über dieses bild erzählen,« 06-07) und sie dieser zugestimmt hat (»BITte«). Erst während ihrer Zustimmung »BITte« begibt sie sich zwei Schritte nach links in das vermeintliche Blickfeld des Roboters (#46.76). Dieses legt zum einen die Überlegung nahe (und hier auf zwei Fallbeispielen basierend), dass die Koordinierungsaktivitäten im Interaktionssystem an das Bearbeiten konkreter Interaktionsaufgaben und sequenzieller Strukturen gebunden sind und nicht notwendigerweise als simple »mechanistische« Reaktion auf Blickabwendungen erfolgen (wie es allerdings bisweilen in der HRI-Literatur suggeriert wird). Zum anderen wird die Orientierung *beider* Teilnehmer – auf der Basis ihrer jeweils unterschiedlichen Handlungslogiken (dynamisch-flexibel lokal interpretierend vs. vorgefertigte Ablaufstrukturen) – an der Beendigung der Eröffnung sichtbar. Nicht nur definiert in diesem Interaktionssystem der Roboter den Übergang von der Eröffnung zum *topic talk*, sondern auch die Besucherin markiert mit dem genauen Zeitpunkt ihres Bearbeitens des visuellen roboterseitigen Koordinierungsangebots ihre Orientierung auf einen solchen Übergang hin. In diesem Sinn machen beide Aktanten gleichzeitig sichtbare Beiträge im Interaktionssystem zur Signalisierung ihrer Bereitschaft zum Übergang zur nächsten Aktivität (wobei sie aufgrund der aktuellen Programmierung faktisch unterschiedlich folgenreich sind).

6.2 Fazit: Ko-Konstruktion – Aufbau von Erwartungen und ihre situativ-emergente Anpassung

Im dritten Fallbeispiel wird Ko-Konstruktion im ›Interaktionssystem Mensch-Roboter‹ aus einer weiteren Perspektive sichtbar, und zwar als Aufbau von nutzerseitigen Erwartungen, wobei diese situativ-emergent im Verlauf der Interaktion angepasst werden. Das responsive *pre-beginning* führt bei der Besucherin zur Annahme, dass der Roboter ein hochkompetenter Interaktionspartner sei und auf sequenzstruktureller Ebene ihre *first turns* bedienen kann. Anders als der Besucher im zweiten Fallbeispiel überträgt die Besucherin hier die Erfahrung der *visuellen* Responsivität im *pre-beginning* auf die *verbale* Handlungsebene. Sie bricht aus den vom Roboter vorgegebenen Ablaufstrukturen aus und übernimmt die Initiative zu *first turns*, womit sie dem Roboter das Ausführen von *second turns* zuweist. Jedoch kann der Roboter die aufgebauten konditionellen Relevanzen nicht bedienen, und die Besucherin reagiert darauf, indem sie dem Roboter (zunächst vorübergehend) die Initiative zu *first turns* überlässt. Hier erfolgt also eine ›Aushandlung‹ zwischen Mensch und Roboter über strukturelle Handlungs- und Anschlussmöglichkeiten im Interaktionssystem. Die zeitversetzte Reaktion der Besucherin auf die Kopfabwendung des Roboters deutet darauf hin, dass Nutzer nicht unbedingt ›automatisch‹ und unmittelbar auf Koordinierungsangebote des Roboters reagieren, sondern dieses als in eine aufgabenbezogene Sequenzstruktur integriert behandeln und damit auf zentrale menschliche Handlungskompetenzen zurückgreifen.

In diesem Fallbeispiel findet die Ko-Konstruktion unter den prinzipiell gleichen Bedingungen wie im zweiten Fallbeispiel statt, die Nutzung der strukturellen Handlungsangebote des Roboters erfolgt anfangs jedoch freier, so dass die Besucherin Schritt für Schritt in die Handlungsmöglichkeiten im Interaktionssystem hinein sozialisiert wird.

7. RÉSUMÉ UND DISKUSSION: KO-KONSTRUKTION IN DER MENSCH-ROBOTER-INTERAKTION

Am Beispiel des Szenarios eines humanoiden Roboters, der als Museums-Guide Besucher einer Kunstaussstellung begrüßt und ihnen Informationen zu verschiedenen Exponaten anbietet, sind wir der Frage nachgegangen, wie Mensch und Roboter – verstanden als *ein* Interaktionssystem (Luhmann 1984) – den Einstieg in ein *focused encounter* organisieren. Der Einstieg ist ein zentraler Moment einer Interaktion, da die Teilnehmer hier nicht nur wechselseitig Zugänglichkeit, sondern auch die Bedingungen und Modalitäten für die anstehende Interaktion herstellen. In der Mensch-Roboter-Interaktion erweist sich dieses als eine besondere Situation, da die Möglichkeiten von Wechselseitigkeit und interaktiver Aushandlung in hohem Maß von den im Robotersystem implementierten Handlungsstrukturen abhängen. In unserem Fall ist der Roboter so ausgestattet, dass traditionell-verbalsprachliche, in ihrem Ablauf vorgefertigte Begrüßungsroutinen (Gruß, Anliegen/Einladung) um ein dynamisch-responsives vorangestelltes *pre-beginning* erweitert werden, in dem die Kopfausrich-

tung des Roboters die Bewegung im Raum eines Besuchers mitvollzieht. Diese Kombination aus initial dynamischen und anschließend festgelegten Handlungsstrukturen erzeugt eine Brechung in der Handlungslogik, so dass in den Studien die Situation eines »Krisenexperiments« (Garfinkel 1967) entsteht.

Für den analytischen Umgang mit solchen Daten erweist sich das von Dausend-schön-Gay/Gülich/Krafft (i. d. B.) vorgestellte Konzept der Ko-Konstruktion als interessant, da es eine analytische Perspektive anbietet, die interaktive Praktiken und Handlungsvollzug sowie *accounts* für die Orientierung der Teilnehmer an ihrer gemeinsamen Aufgabenbearbeitung fokussiert. Attraktiv ist dabei für die HRI ganz besonders, dass nicht die Frage im Vordergrund steht, ob Ko-Konstruktion stattfindet, sondern *wie* sie von den Teilnehmern gestaltet wird. Dieses erweist sich für unser Erkenntnisinteresse als Vorteil gegenüber der in der Technik-Soziologie diskutierten Frage, ob der Begriff der Interaktion angemessen für Situationen der Begegnung zwischen Mensch und Maschine sei. Hierfür werden in jüngerer Zeit z. B. die Konzeption einer »hybriden sozio-technischen Konstellation« (Rammert/Schulz-Schaeffer 2002) und eines gradualisierten Handlungsbegriffs als Alternativen angeboten, die die »Intelligenz« eines technischen Systems nicht als individuelle Leistung verstehen, sondern als eine sozio-technische Größe, die über die Zuschreibungsprozesse der Teilnehmer empirisch erforscht werden kann. Der Begriff der Ko-Konstruktion bietet sich demgegenüber als analytische Perspektive an, die den Fokus unmittelbar auf die interaktive Herstellungsleistung und die gemeinsamen Praktiken der Teilnehmer legt. Für die HRI scheint er auch deshalb gut geeignet zu sein, da besonders in der »Krisenexperimenthaftigkeit« von Studien mit frühen Forschungsprototypen explizite und implizite *accounts* für Handlungen bzw. das Ausbleiben von erwarteten Aktionen entstehen.

Mit diesem Fokus und konzeptuellen Hintergrund hat die empirische Analyse von drei Fallbeispielen gezeigt:

(1) Auch wenn an der Interaktionsoberfläche in der Eröffnung die verbalsprachlichen Handlungsroutinen über das Aufbauen starker konditioneller Relevanzen die Besucher dazu bringen, in das Abarbeiten der Routine einzusteigen, sind damit noch nicht notwendigerweise die tatsächlichen interaktiven Aufgaben mit ihren Funktionalitäten des Zugänglichkeitsherstellens erbracht. Für den Nutzer wird damit nicht ersichtlich, was im Interaktionssystem als relevante Anschlusshandlung gilt.

(2) Demgegenüber steht ein Einstieg mit responsivem *pre-beginning*, in dem die Besucher den Roboter als ein dynamisch-responsives Gegenüber erleben, das u. U. Wechselseitigkeit von Wahrnehmung signalisiert und Positionsveränderungen des Besuchers mitverfolgt. Hierbei werden andere Erwartungen für relevante Anschluss-handlungen aufgebaut, die sich im unmittelbaren Verwenden von natürlichsprachlichen kommunikativen Ressourcen manifestiert. Inwiefern bzw. unter welchen Bedingungen sich die Beobachtungen aus (1) und (2) über die betrachteten Fallbeispiele hinaus als ein systematisch reproduzierbarer Zusammenhang erweisen, wird künftig auf größerer Datenbasis zu untersuchen sein.

(3) Interessant ist insbesondere der Vergleich von strukturell ähnlichen Momenten in den drei Fallbeispielen, an denen der Roboter den Kopf vom aktuellen Interaktionspartner abwendet, und zwar in eine Richtung, in der sich ein weiterer Besucher befindet. In diesen Situationen lässt sich feststellen: (i) Beim Einstieg *ohne* Erfahrung

von Responsivität behandelt der Nutzer diese Veränderung in der Kopforientierung des Roboters als eine Abwahl seiner Person als primären Interaktionspartner und steigt bald darauf aus der Interaktion aus. (ii) Beim Einstieg *mit* Erfahrung von Responsivität folgt der Nutzer kurz der neuen Aufmerksamkeitsausrichtung und repositioniert sich dann wieder im (vermeintlichen) Blickfeld des Roboters. Damit manifestiert er aktiv den Anspruch, seinen bisherigen Beteiligungsstatus aufrecht zu erhalten. Wichtig ist hierbei insbesondere die sequenzielle Platzierung dieser Ereignisse: Die Besucher behandeln die Veränderung der Kopfausrichtung erst dann als interaktionsrelevant, wenn sie selbst in den angebotenen nächsten Handlungsschritt einsteigen. Solche roboterseitigen Verhaltensweisen bieten also (in zukünftiger Forschung weiter zu explorierende) Möglichkeiten für das Interaktionsmanagement, funktionieren aber kaum als situationsunabhängige Automatismen. Über die Rekonfiguration von Interaktionsraum – wie sie auch in menschlicher Interaktion an dieser Stelle erfolgt – markieren die Nutzer ihre Orientierung an der bevorstehenden Beendigung der Eröffnung und ihre Bereitschaft zum Übergang in den *topic talk*.

(4) In den beiden Fällen des Einstiegs mit *pre-beginning* zeigt sich darüber hinaus, dass die erfahrene Responsivität des Roboters auf verschiedenen Ebenen interpretierbar ist: auf visueller vs. auf verbalsprachlicher Ebene. Die Unterstellung der Besucherin im dritten Fallbeispiel, dass der Roboter Interaktionskompetenz für die Übernahme von *second turns* habe, verweist auf die Notwendigkeit von detaillierteren Studien zum Zusammenhang zwischen Roboterverhalten und darauf basierenden nutzerseitigen Zuschreibungsleistungen beim Einstieg in ein *focused encounter*.

Insgesamt ergeben sich aus diesen Beobachtungen zur Ko-Konstruktion in der Mensch-Roboter-Interaktion Implikationen in zweierlei Richtung:

Interaktionsforschung: Das Konzept der Ko-Konstruktion erweist sich als fruchtbar und anschlussfähig für die Beschäftigung mit Situationen der Mensch-Roboter-Interaktion. Es erscheint kompatibel mit unserem Ansatz, die HRI als ein Instrument für die Erforschung von Interaktion zu verstehen und dabei zum einen aus der menschlichen Interaktion bekannte Verfahren auf ihre Grenzen und Implikationen in der HRI zu erproben und in iterativem Design für die Modellierung auf die spezifisch technische Konstellation hin zu verändern, und zum anderen es unter der Perspektive von »Krisenexperimenten« zu betrachten, die die bekannten Interaktionsstrukturen in spezifischer – und vom Forscher modellierbarer – Weise aufbrechen, um damit einen analytischen Zugriff auf einzelne interaktionsrelevante Aspekte zu erhalten. Für die Frage des *interactional achievement* liefert die Beschäftigung mit HRI den Hinweis auf eine notwendige Unterscheidung zwischen interaktiven Oberflächen-Strukturen und den zugrundeliegenden Handlungslogiken. Diese sind in HHI häufig schwer zu fassen, in HRI sind sie Teil des (vom Forscher beeinflussbaren) Interaktionsdesigns.

Mensch-Roboter-Interaktion: Für die Gestaltung des Interaktionseinstiegs und das roboterseitige Eröffnungsverhalten ist ein dynamischerer Zugang als bisher praktiziert möglich. Bei dessen weiterer Exploration wird genauer zu untersuchen sein, durch welche Verhaltensweisen und interaktiven Verfahren in der Eröffnung Roboter Hinweise auf ihre Interaktionskompetenzen und die tatsächlich relevanten Anschlusshandlungen im Interaktionssystem liefern können. Damit entstehen – in (künftig näher zu definierenden) Grenzen – Möglichkeiten, um mit der grundlegenden

Unvorhersehbarkeit und Kontingenz von Interaktion beim Design von Interaktionsverhalten für Roboter etwas besser umgehen zu können. Es stellt sich die Frage, wie auf dieser Basis das Prinzip von projektiven Strukturen umgesetzt werden kann. Insgesamt erweist sich der Ansatz, Mensch und Roboter als ein Interaktionssystem zu verstehen, als fruchtbar und geht konzeptionell über die *human in the loop*-Ansätze hinaus. Hier eröffnet sich für die Gestaltung von Interaktionsfähigkeiten die Frage, welche Handlungsmöglichkeiten das System dem Nutzer für die Gestaltung von *interactional achievement* und damit für das Einbringen seiner Interaktionskompetenzen (die für das Funktionieren von HRI notwendig sind ((Pitsch, accepted)) bietet.

LITERATUR

- Auer, Peter (2009): »Projection and minimalistic syntax in interaction«, in : *Discourse Processes* 46(2), S. 180-205.
- Bergmann, Jörg (1988): »Haustiere als kommunikative Ressourcen.«, in: Soeffner, Hans-Georg (Hg.): *Kultur und Alltag*, Göttingen: Schwartz, S. 299-312.
- Bohus, Dan/Horvitz, Eric (2009): »Learning to predict engagement with a spoken dialog system in open-world settings«, in: *Proceedings of the SIGDIAL 2009 conference*, S. 244-252.
- Breazeal, Cynthia (2003) : »Toward sociable robots«, in : *Robotics and Autonomous Systems* 42(3-4), S. 167-175.
- Button, Graham (1990): »Going Up a Blind Alley. Conflating Conversation Analysis and Computational Modeling«, in: Paul Luff/Nigel Gilbert/David M. Frohlich (Hg.), *Computers and conversation*, London, San Diego: Academic Press Limited, S. 67-90.
- Dausendschön-Gay, Ulrich/Krafft, Ulrich (2009): »Preparing next actions in routine activities«, in: *Discourse Processes* 46(2), S. 247-268.
- Fischer, Kerstin (2006) : *What computer talk is and isn't: Human-Computer conversation as intercultural communication*, Weinberg: AQ-Verlag.
- Frohlich, David/Luff, Paul (1990): »Applying the technology of conversation to the technology for conversation«, in: Paul Luff/Nigel Gilbert/David M. Frohlich (Hg.), *Computers and conversation*, London, San Diego: Academic Press Limited, S. 187-220.
- Garfinkel, Harold (1967): *Studies in ethnomethodology*, Englewood Cliffs, NJ: Prentice-Hall.
- Heath, Christian (1992) : »Gesture's discreet tasks. Multiple relevancies in visual conduct and in the contextualisation of language«, in : Peter Auer/Aldo Di Luzio (Hg.), *The contextualization of language*, Amsterdam: John Benjamins, S. 101-133.
- Hegel, Frank/Lohse, Manja/Wrede, Britta (2009): »Effects of visual appearance on the attribution of applications in social robotics«, in: *RO-MAN 2009*, S. 64-71.
- Holthaus, Patrick/Pitsch, Karola/Wachsmuth, Sven (2011): »How can I help?«, *International Journal of Social Robotics* 3(4), S. 383-393.

- Kanda, Takayuki/Ishiguro, Hiroshi (2012): *Human-Robot interaction in social robotics*, Boca Raton: CRC Press.
- Kendon, Adam/Ferber, Andrew (1973) : »A description of some human greetings«, in: Richard Phillip Michael/John Hurrell Crook (Hg.), *Comparative ecology and behavior of primates*, London, New York: Academic Press, S. 591-668.
- Krummheuer, Antonia (2010): *Interaktion mit Virtuellen Agenten? Zur Aneignung eines ungewohnten Artefakts*, Stuttgart: Lucius & Lucius.
- Levinson, Stephen C. (2006): »On the human "interaction engine"«, in: In Nick J. Enfield/Stephen C. Levinson (Hg.), *Roots of human sociality. Culture, cognition and interaction*, Oxford, New York: Berg, S. 39-69.
- Lison, Pierre (2015, in press): *A hybrid approach to dialogue management based on probabilistic rules*. *Computer Speech & Language*.
- Lohse, Manja (2009): *Investigating the influence of situations and expectations on user behavior - empirical analyses in human-robot interaction*. PhD Thesis, Bielefeld University.
- Luhmann, Niklas (1984): *Soziale Systeme*, Frankfurt am Main: Suhrkamp.
- Mondada, Lorenza (2009): »Emergent focused interactions in public places: A systematic analysis of the multimodal achievement of a common interactional space«, in: *Journal of Pragmatics* 41(10), S. 1977-1997.
- Mortensen, Kristian/Hazel, Spencer (2014): »Moving into interaction - social practices for initiating encounters at a help desk«, in: *Journal of Pragmatics* 62, S. 46-67.
- Pitsch, Karola/Kuzuoka, Hideaki/Suzuki, Yuya/Süssenbach, Luise/Luff, Paul/Heath, Christian (2009): »"The first five seconds". Contingent stepwise entry into an interaction as a means to secure sustained engagement«, in: *RO-MAN 2009*, S. 985-991.
- Pitsch, Karola/Lohan, Katrin S./Rohlfing, Katharina/Saunders, Joe/Nehaniv, Christopher L./Wrede, Britta (2012): »Better be reactive at the beginning. Implications of the first seconds of an encounter for the tutoring style in human-robot-interaction«, in: *Ro-Man 2012*, S. 974-981.
- Pitsch, Karola/Vollmer, Anna-Lisa/ Mühlig, M. (2013): »Robot feedback shapes the tutor's presentation. How a robot's online gaze strategies lead to micro-adaptation of the human's conduct«, in: *Interaction Studies* 14(2), S. 268-296.
- Pitsch, Karola/Vollmer, Anna-Lisa/Rohlfing, Katharina/Fritsch, Jannik/Wrede, Britta (2014): »Tutoring in adult-child-interaction: On the loop of the tutor's action modification and the recipient's gaze«, in: *Interaction Studies* 15(1), S. 55-98.
- Pitsch, Karola/Wrede, Sebastian (2014): »When a robot orients visitors to an exhibit. Referential practices and interactional dynamics in real world HRI«, in: *Ro-Man 2014*, S. 36-42.
- Pitsch, Karola (accepted): »Limits and opportunities for mathematizing communicational conduct for social robotics in the real-world? - Towards enabling a robot to make use of the human's competences«, in: Gesa Lindemann (Hg.), *Going beyond the laboratory* (provisional title).
- Rammert, Werner/Schulz-Schaeffer, Ingo (2002): »Technik und Handeln. Wenn soziales Handeln sich auf menschliches Verhalten und technische Abläufe verteilt«,

- in: Können Maschinen handeln? Soziologische Beiträge zum Verhältnis von Mensch und Technik, Frankfurt/Main: Campus, S. 11-64.
- Schegloff, Emanuel A. (1986): »The routine as achievement«, in: *Human Studies* 9, S. 111-151.
- Schegloff, Emanuel A. (1992): »To Searle on conversation: A note in return«, in: John R. Searle (Hg.), (On) Searle on conversation, Amsterdam, Philadelphia: Benjamins, S. 113-128.
- Schegloff, Emanuel A. (1996): »Issues of relevance for discourse analysis: Contingency in action, interaction, and co-participant context«, in: Eduard H. Hovy/Donia R. Scott (Hg.), *Computational and conversational discourse: Burning issues - an interdisciplinary account*, Berlin: Springer, S. 3-38.
- Searle, John R. (1992): »Conversation«, in: John R. Searle/Herman Parret/Jef Verschueren (Hg.), (On) Searle on conversation, Amsterdam: Benjamins, S. 7-29.
- Shiomi, Mashiro/Kanda, Takayuki/Ishiguro, Hiroshi/Hagita, Norihiro (2010): »A larger audience, please!: Encouraging people to listen to a guide robot«, in: *HRI '10: Proceeding of the 5th ACM/IEEE international conference on human-robot interaction*, S. 31-38.
- Skantze, Gabriel/Schlangen, David (2009): »Incremental dialogue processing in a micro-domain«, in: *EACL 2009*, S. 745-753.
- Streeck, Jürgen/Jordan, J. Scott (2009): »Projection and anticipation: The forward-looking nature of embodied communication«, in: *Discourse Processes* 46(2), S. 93-102.
- Suchman, Lucy A. (1987): *Plans and situated actions. The problem of human-machine communication*, Cambridge: Cambridge University Press.

DuEPublico

Duisburg-Essen Publications online

UNIVERSITÄT
DUISBURG
ESSEN

Offen im Denken

ub

universitäts
bibliothek

Dieser Text wird über DuEPublico, dem Dokumenten- und Publikationsserver der Universität Duisburg-Essen, zur Verfügung gestellt. Die hier veröffentlichte Version der E-Publikation kann von einer eventuell ebenfalls veröffentlichten Verlagsversion abweichen.

DOI: 10.14361/9783839432952-013

URN: urn:nbn:de:hbz:464-20201119-143936-3

Pitsch, Karola: Ko-Konstruktion in der Mensch-Roboter-Interaktion: Kontingenz, Erwartungen und Routinen in der Eröffnung.

In: *Ko-Konstruktionen in der Interaktion. Die gemeinsame Arbeit an Äußerungen und anderen sozialen Ereignissen / Gülich, Elisabeth; Krafft, Ulrich; Dausendschön-Gay, Ulrich (Hrsg.). Bielefeld : transcript Verlag, 2015, S. 229 - 258.*

eISBN: 978-3-8394-3295-2 - DOI: <https://doi.org/10.14361/9783839432952>

© transcript Verlag (2015) Alle Rechte vorbehalten.