

### Von Menschen und Maschinen: Interdisziplinäre Perspektiven auf das Verhältnis von Gesellschaft und Technik in Vergangenheit, Gegenwart und Zukunft ; Proceedings der 3. Tagung des Nachwuchsnetzwerks "INSIST", 05.-07. Oktober 2018, Karlsruhe

Ahner, Helen (Ed.); Metzger, Max (Ed.); Nolte, Mathis (Ed.)

Erstveröffentlichung / Primary Publication

Sammelwerk / collection

#### Empfohlene Zitierung / Suggested Citation:

Ahner, H., Metzger, M., & Nolte, M. (Hrsg.). (2020). *Von Menschen und Maschinen: Interdisziplinäre Perspektiven auf das Verhältnis von Gesellschaft und Technik in Vergangenheit, Gegenwart und Zukunft ; Proceedings der 3. Tagung des Nachwuchsnetzwerks "INSIST", 05.-07. Oktober 2018, Karlsruhe* (INSIST-Proceedings, 3). <https://nbn-resolving.org/urn:nbn:de:0168-ssoar-67663-1>

#### Nutzungsbedingungen:

Dieser Text wird unter einer CC BY-NC-ND Lizenz (Namensnennung-Nicht-kommerziell-Keine Bearbeitung) zur Verfügung gestellt. Nähere Auskünfte zu den CC-Lizenzen finden Sie hier:

<https://creativecommons.org/licenses/by-nc-nd/3.0/deed.de>

#### Terms of use:

This document is made available under a CC BY-NC-ND Licence (Attribution-Non Commercial-NoDerivatives). For more information see:

<https://creativecommons.org/licenses/by-nc-nd/3.0>

---

INSIST-Proceedings  
Julia Engelschalt, Arne Maibaum (Hrsg.)

---

# Von Menschen und Maschinen

Interdisziplinäre Perspektiven auf  
das Verhältnis von Gesellschaft  
und Technik in Vergangenheit,  
Gegenwart und Zukunft

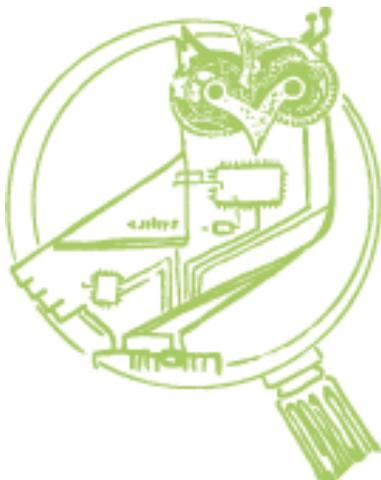
---

Proceedings der 3. Tagung  
des Nachwuchsnetzwerks „INSIST“  
05.-07. Oktober 2018, Karlsruhe

---

Band 3  
Herausgegeben von  
Helen Ahner, Max Metzger & Mathis Nolte

---



# Inhaltsverzeichnis

Geleitwort.....	i
Editorische Notiz.....	iii
Fährt selbst und ständig: Empirische Nutzeranalysen eines automatisierten Mobilitätsangebotes an einem Großklinikum und im ÖPNV.....	1
Lina Kluy, Stefan Blüher und Jan C. Zöllick	
Master or Servant? Der Wandel im Mensch-Maschine-Verhältnis in der internationalen zivilen Luftfahrt des 20. Jahrhunderts.....	15
Sabrina Lausen	
Wo ist der Mensch in der automatisierten Produktion? Eine aktuelle Frage aus historischer Perspektive.....	35
Nikolai Ingenerf, Moritz Müller und Nora Thorade	
Neue Wege des Passing mit Prothesen? Zur Kosmetisierung der Beinprothetik in der Bundesrepublik Deutschland der 1960er und 1970er Jahre.....	55
Mathis Nolte	
Somatisch-visuelle Aushandlung embryonalen Lebens. Zur Konstitution embryonaler Wesen am Beispiel Exitus im Uterus.....	77
Nico Wettmann	
Der Mensch als zwecklose Maschine? Descartes' Philosophie in der Kontroverse .....	93
Daniel Neumann	
Künstliche Intelligenz in der Science-Fiction: Mehr Magie als Technik.....	105
Isabella Hermann	
Science fiction is what got me into the field. Elemente der Popkultur als Vermittlungsstrategien im Diskurs um künstliche Intelligenz.....	119
Rebecca Bachmann	
Cyborg als Metapher. Haraway mit Blumenberg lesen.....	141
Lisa Schurrer	
Humanoide Roboter und virtuelle Agenten als Kommunikationsteilnehmer? Konversationsanalytische Studien der Mensch-Maschine-Interaktion.....	159
Indra Bock und Henning Mayer	
„Naturgetreu jedoch beschleunigt“ - Wie im Projektionsplanetarium Maschinen die Weltdeutung übernahmen.....	183
Helen Ahner	

Träumen rote KIs von Lenin? Die kybernetische Hypothese zwischen Cybersyn, Kapitalismus und anarchistischer Politik.....	203
David Kipscholl und Alexander Kurunzi	
„Die Maschine hat den Piloten abgeworfen“ - Horkheimers Kritik der instrumentellen Vernunft metaphorologisch gelesen.....	223
Andreas Brenneis	
Autor*innenverzeichnis.....	237

# Geleitwort

Das Interdisciplinary Network for Studies Investigating Science and Technology (INSIST) ist eine disziplin- und standortübergreifende Initiative des Nachwuchses im Feld der Wissenschafts- und Technikforschung. INSIST richtet sich an Nachwuchswissenschaftler\*innen, Studierende und alle Interessierten, die sich für Fragen der Wissenschafts- und Technikforschung begeistern und nach Möglichkeiten des thematischen wie auch informellen Erfahrungsaustausch suchen. Gegründet wurde das Netzwerk im Oktober 2013 in Bielefeld.

Die selbstgewählten Ziele der Förderung und Vernetzung des Nachwuchses sind weder an spezifische akademische Einrichtungen noch an Zugehörigkeiten zu bestimmten akademischen Disziplinen gebunden. INSIST versteht sich als Plattform zur Erhöhung der inneren und äußeren Sichtbarkeit von in der Wissenschafts- und Technikforschung meist eher unterrepräsentierten Gruppen. Das Netzwerk beschränkt sich in seinen Aktivitäten daher nicht ausschließlich auf klassische akademische Nachwuchsgruppen wie Postdocs und Doktorand\*innen, sondern bezieht in seine Veranstaltungen bewusst auch Studierende und andere Interessierte mit ein.

Dem Motto „Vom Nachwuchs für den Nachwuchs“ folgend, hat es sich INSIST unter anderem zur Aufgabe gemacht, alle zwei Jahre an wechselnden Standorten eine interdisziplinäre Nachwuchstagung zu organisieren. Diese sollen Nachwuchswissenschaftler\*innen einen vergleichsweise geschützten Raum bieten, erste Erfahrungen mit eigenen wissenschaftlichen Vorträgen und Workshops zu sammeln.

Der vorliegende Proceedings-Band ist aus der dritten INSIST-Nachwuchstagung „Von Menschen und Maschinen. Interdisziplinäre Perspektiven auf das Verhältnis von Gesellschaft und Technik in Vergangenheit, Gegenwart und Zukunft“, hervorgegangen, die vom 05. bis 07. Oktober 2018 am Karlsruher Institut für Technologie stattfand.

Über 50 teilnehmende Wissenschaftler\*innen haben im Rahmen von 30 Vorträgen und 3 Workshops die Bestimmungen, Aneignungen und Verhältnisnahmen von Mensch und Maschine reflektiert und diskutiert.

Die Keynotes von Martina Heßler zum Thema „*Mensch|Maschinen. Perspektiven einer historischen Technikanthropologie*“ und Gabriele Gramelsberger zum Thema „*Parallelgesellschaft der Maschinen. Wie weit geht die Automatisierung?*“ bereicherten das Abendprogramm mit synoptischen Betrachtungen und theoretischen Ausblicken um einen gesellschaftspolitischen und forschungspragmatischen Blick aufs Tagungsthema.

INSIST begrüßt und unterstützt die Forschung und Präsentation in neuen und innovativen Formaten. Mit der interaktiven Kunstinstallation „*Nachrichten an mich*“ von Maja Urbanczyk, die audiovisuelle Live-Performance „*We all learn to desire the same things. Allowing images to become a tool*“ der Gruppe Ilaria Atonali, sowie einer Podiumsdiskussion zu wissenschaftlichen Publikationspro-

zessen mit Vertreter\*innen der Zeitschriften *Technikgeschichte*, *NTM* (Zeitschrift für Geschichte der Wissenschaft, Medizin und Technik), *TATuP* (Zeitschrift für Technikfolgenabschätzung in Theorie und Praxis) und *NanoEthics. Studies of New and Emergig Technologies* konnten wir die Reflektion des Tagungsthemas jenseits traditioneller Vortragsformate sehr produktiv einbinden.

Das vollständige Tagungsprogramm kann auf der INSIST Website unter <http://insist-network.com/insist-tagung-2018-programm/> eingesehen werden.

Wir bedanken uns an dieser Stelle noch einmal sehr herzlich bei allen Tagungsteilnehmer\*innen für ihre jeweiligen Anregungen und Diskussionsbeiträge. Unsere Dankbarkeit gilt auch dem Institut für Technikzukünfte, dem daran angeschlossenen Teilinstitut für Technikgeschichte sowie dem Institut für Germanistik am Karlsruher Institut für Technologie, ohne deren großzügige finanzielle und organisatorische Unterstützung, die Durchführung der Tagung nicht möglich gewesen wäre.

Nicht zuletzt gilt unser besonderer Dank den Herausgeber\*innen dieses Bandes, die mit ihrem Engagement dem Anliegen von INSIST, jungen Forscher\*innen der Wissenschafts- und Technikforschung eine Plattform und eine eigenständige Stimme zu geben, hervorragende Form und Umsetzung gegeben haben.

Helen Ahner & Franz Kather  
Sprecher\*innen von INSIST

# Editorische Notiz

Der vorliegende dritte Band der INSIST-Proceedings-Reihe versammelt 12 zur Veröffentlichung ausgearbeitete Beiträge der dritten INSIST-Nachwuchstagung „Von Menschen und Maschinen. Interdisziplinäre Perspektiven auf das Verhältnis von Gesellschaft und Technik in Vergangenheit, Gegenwart und Zukunft“.

Die Reihenfolge der hier zusammengestellten Texte reproduziert weder die zeitliche Abfolge im Tagungsprogramm, noch soll durch die gewählte Anordnung eine qualitative Wertung vorgenommen werden.

Um der disziplinären Vielfalt und Unterschiedlichkeit der einzelnen Beiträge gerecht zu werden, haben wir uns – wie auch schon bei den vorangehenden Proceedings Bänden – dazu entschieden, die Wahl von Zitiersystemen in Fußnoten und bibliographischen Angaben unseren Autor\*innen zu überlassen und die jeweiligen Texte lediglich im Layout zu vereinheitlichen. Auch der Umgang mit geschlechtergerechten Schreibweise blieb den Autor\*innen überlassen.

Alle Beiträge haben ein doppelt anonymisiertes Peer-Review-Verfahren durchlaufen. Wir möchten uns daher nicht nur bei den Autor\*innen für ihre Einreichungen, sondern auch bei den jeweiligen Gutachter\*innen für ihre konstruktiven Anmerkungen und Verbesserungsvorschläge bedanken.

Für die erneute Möglichkeit der Online-Publikation im Social Science Open Access Repository (SSOAR) danken wir zudem dem GESIS Leibnitz-Institut für Sozialforschung.

Julia Engelschalt, Universität Bielefeld  
Arne Maibaum, Technische Universität Berlin  
Helen Ahner, Universität Tübingen  
Max Metzger, Technische Universität Dresden  
Matthis Nolte, Stadtarchiv Löhne

# Humanoide Roboter und virtuelle Agenten als Kommunikationsteilnehmer?

## Konversationsanalytische Studien der Mensch-Maschine-Interaktion

---

Indra Bock und Henning Mayer

### 1. Einleitung

Die Vorstellung menschenähnlicher Interaktion mit technischen Artefakten wird zunehmend real und damit sozialwissenschaftlich empirisch relevant. Besonders soziale Roboter und verkörperte Agenten, die als Humanoide konzipiert sind, fordern die Soziologie heraus (Muhle 2018a), denn diese sollen in der Lage sein, in soziale Beziehungen mit Menschen zu treten (Breazeal 2002) und mittels natürlicher Sprache mit ihnen zu interagieren. Entsprechende Systeme sollen zukünftig in den Alltag (z.B. im Pflegebereich) integriert werden und dort nicht nur als Werkzeuge sondern vielmehr als Kommunikationspartner von Menschen fungieren (Zhao 2009; Meister 2011). Heute können Menschen humanoiden Robotern z.B. schon im Heinz-Nixdorf-Museumsforum (HNF) in Paderborn, dem weltgrößten Computermuseum<sup>1</sup>, als Teil der Ausstellung gegenüber treten und Kontakt zu ihnen aufnehmen. Die Frage nach den Formen und Grenzen der Kommunikation zwischen Mensch und Technik stellt sich also bereits aktuell.

Vor diesem Hintergrund beschäftigt sich unser Beitrag empirisch mit Begegnungen zwischen Besucher\*innen des HNF mit zwei Humanoiden in der Ausstellung, dem Roboter Nadine und dem Agenten MAX, um die Charakteristika der ‚Interaktionen‘<sup>2</sup> herauszuarbeiten. Hierbei nehmen wir eine kommunikationstheoretisch fundierte Perspektive ein. D.h. im Fokus steht (in erster Linie) das kommunikative Geschehen *zwischen* Mensch und Maschine und nicht etwa die Frage nach subjektiven Deutungen des technischen Gegenübers durch Men-

---

<sup>1</sup> Unsere Daten wurden im Rahmen des von der DFG geförderten Forschungsprojektes ‚Kommunikation an den Grenzen der Sozialität‘ ([www.gepris.dfg.de/gepris/projekt/318924833](http://www.gepris.dfg.de/gepris/projekt/318924833), Abfrage: 7.6.2018) erhoben. Möglich wurden die Aufzeichnungen dank der großartigen Unterstützung durch das HNF und hier insbesondere Dr. Stefan Stein.

<sup>2</sup> Wir verwenden den Begriff ‚Interaktion‘ in Anführungszeichen, da fraglich ist, ob und wie sich Anwesenheit und Kommunikation (vgl. Kieserling 1999) in Mensch-Roboter-Begegnungen etablieren.

schen (Muhle 2013; Muhle 2018a). Diese Perspektive lässt sich methodisch optimal mit der ethnomethodologischen Konversationsanalyse (KA) einlösen, die sequenzanalytisch die sinnhafte Verkettung der kommunikativen Ereignisse rekonstruiert (Muhle 2016; Schneider 2008). Die KA hat sich spätestens seit Lucy Suchmans (1987) richtungsweisender Studie als Methode zur Erforschung der Mensch-Maschine-Interaktion etabliert. Mit den Workplace Studies ist eine Forschungstradition entstanden, die dem situativen und interaktiven Gebrauch von Technik empirisch nachgeht und auch im noch jungen Bereich der Human-Robot-Interaction (HRI) hat sich die KA (bzw. Ethnomethodologie) bereits bewährt (Muhle 2018b, Alač 2016, Pitsch 2015, Krummheuer 2010).

Konkret möchten wir in unserem Beitrag zwei Fallbeispiele vorstellen, anhand derer wir mit unterschiedlichen analytischen Schwerpunkten und Erhebungsmethoden nicht nur spezifische Besonderheiten der jeweiligen Begegnungen herausarbeiten, sondern auch zeigen, wie das konversationsanalytische Vorgehen ganz im Sinne der methodologischen Forderung des „unique adequacy requirement of methods“ (Garfinkel/Wieder 1992) den jeweiligen Gegenständen angepasst werden kann. So werden in die Analyse der Begegnung mit Nadine auch die nonverbalen Modalitäten (Blick, Gestik, Mimik, Körperbewegungen) gemäß einer multimodalen Interaktionsanalyse (Schmitt 2015) einbezogen, um auf diese Weise zu zeigen, wie in der Begegnung sprachliche und körperliche Aktivitäten auseinander driften, während in die Analyse der Begegnung mit MAX auch die Innenwelt des Agenten einbezogen wird, um zu zeigen, wie das System operiert und um dessen teilweise problematische Anschlüsse *accountable* zu machen.<sup>3</sup> Die Verbindung dieser beiden Perspektiven erlaubt u.E. die Beteiligung adressierbarer technischer Systeme an sozialer Kommunikation als ein vielfach zu synchronisierendes Arrangement in den Blick zu nehmen, in dem sowohl unterschiedliche Modalitäten der Kommunikation als auch verschiedene Formen der kommunikativen Aneignung von (innenweltlichen) Entkopplungen des technischen Schlussfolgerns aufeinandertreffen.

## **2. Nadine: Entkopplung von Sprache und Körper**

### *2.1 Der soziale Roboter Nadine*

Der Roboter Nadine ist seiner Entwicklerin Nadia Magnenat Thalmann optisch nachempfunden und sieht zum Verwechseln menschlich aus. Als „sitting pose robot“ (Magnenat Thalmann et al. 2017) kann Nadine Oberkörper, Arme und Kopf samt Mund, Brauen und Lidern in 27 Freiheitsgraden bewegen und ist in der Lage rudimentäre Mimik und Gestik zu zeigen. In der Ausstellung im HNF Paderborn ist sie hinter einem Schreibtisch mit Laptop platziert und erhält damit einen augenscheinlichen Selbstbezug. Über ein Kinect-Kamerasystem registriert Nadine Menschen in ihrem Wahrnehmungsbereich und reagiert auf sie, indem sie von ihrer Beschäftigung aufschaut, Blickkontakt zu näherkommenden Besucher\*innen aufnimmt (s. Abb. 1) sowie lächelt und winkt.

---

<sup>3</sup> Für den Zugang zum Programmcode danken wir Prof. Dr. Ipke Wachsmuth, dem ehem. Leiter der AG Wissensbasierte Systeme an der Universität Bielefeld und Dr. Thies Pfeiffer.

Die Besucher\*innen können über ein Mikrofon, welches dezent vor dem Schreibtisch an einem Ständer angebracht ist, mit dem Roboter sprechen. Obwohl Nadine also leibhaftig vor den Besucher\*innen sitzt, müssen diese doch technisch vermittelt mit ihr Kontakt aufnehmen. Nadines Antwort wird durch versteckte Lautsprecher wiedergegeben – ihr Mund bewegt sich dazu synchron. Nadine besitzt also zwar Augen, Ohren und Lippen; ihre Wahrnehmungs- und Kommunikationsorgane sind aber weitgehend auf technische Geräte um sie herum distribuiert. Dennoch stellt ihr Körper die zentrale Schnittstelle für den Kontakt mit den Museumsbesucher\*innen dar und weckt Erwartungen hinsichtlich der Möglichkeit, sich mit ihr ganz natürlich zu verständigen.<sup>4</sup> Diese werden jedoch in verschiedenen Hinsichten gebrochen, wie im Folgenden mithilfe einer multimodalen Interaktionsanalyse deutlich wird.



Abb. 1: Nadine nimmt Blickkontakt auf.

## 2.2 Multimodale Interaktionsanalyse

Die Multimodale Interaktionsanalyse stellt gewissermaßen keine Erweiterung der klassischen Konversationsanalyse dar, sondern die klassische KA nimmt umgekehrt eine Reduktion auf oder zumindest Priorisierung von Verbalität vor (Schmitt 2015: 43)<sup>5</sup>. Diese Priorisierung wird in der multimodalen Interaktionsanalyse aufgehoben zugunsten einer Egalität aller beobachtbaren Ausdrucksmodalitäten. Das bedeutet in erster Linie, dass die sequenzielle Analyse um die Dimension der Simultanität erweitert wird. Ein kommunikativer Beitrag wird nicht nur in seiner Verkettung mit dem sequenziell vorausgehenden oder fol-

<sup>4</sup> Analysen des Datenmaterials zeigen, dass die Besucher\*innen den Roboter häufig unmittelbar ansprechen und das Mikro erst nach ausbleibender Antwort bemerken. Auch richtet sich der Blick der Personen in Erwartung einer Antwort von Nadine auf das Gesicht des Roboters.

<sup>5</sup> Die Workplace Studies sind hier ausgenommen; ihr Fokus lag schon immer und ganz gezielt auch auf nonverbalen Elementen des Handlungszusammenhangs wie verkörpertem Wissen oder der Manipulation von Gegenständen und Technik (Knoblauch 2000). Ihre Forschungsgegenstände sind jedoch eher Computer unterstütztes kooperatives Arbeiten oder Human-Computer-Interaction (HCI), bei denen der Fokus auf den beteiligten Menschen liegt.

genden Beitrag interpretiert, sondern auch im Zusammenspiel mit den gleichzeitig stattfindenden kommunikativen Elementen über verschiedene Modalitäten hinweg. Insbesondere geht es hierbei um Blickrichtung, Körperbewegung und -postur, Konstellation der Beteiligten zueinander, Mimik, Gestik und Manipulation von Objekten sowie natürlich Sprache und Intonation (Mondada/Schmitt 2010). Wie in der klassischen KA auch werden Bedeutungs- und Erwartungshorizonte kommunikativer Beiträge aufgefächert und durch die Kontrastierung mit anschließenden oder simultanen Beiträgen geschlossen. Ziel der Analyse ist es, interaktive Ordnungsstrukturen und Muster zu rekonstruieren und so Formen von Kommunikation oder reziproken Aktivitäten zu beschreiben. Was für Menschen unbewusst und selbstverständlich gelingt: eine kommunikative Handlung multimodal hervorzubringen, ist für humanoide Roboter eine große Herausforderung (Magenat Thalmann et al. 2016). Wie die folgende Analyse zeigt, hat auch Nadine Schwierigkeiten, Körper und Verbalität im Einklang zu halten.

Die folgende Begegnung zwischen Museumsgästen und Nadine wurde im Rahmen eines videographischen Forschungssettings (Knoblauch 2013) mit zwei Kameras aufgezeichnet und multimodal transkribiert. Die Darstellung des Transkriptes erfolgt in Tabellenform als Partitur. Die notierten Ereignisse sind entsprechend diachron von links nach rechts und synchron von oben nach unten lesbar. In der oberen Zeile sind Nummer und Zeitmarke<sup>6</sup> der jeweiligen Notation vermerkt. An relevanten Stellen wurde die Pausen-Dauer (von der letzten verbalen Aktivität ausgehend bis zum jeweiligen Ereignis) eingetragen.

### *2.3 Fallbeispiel Nadine*

Der Roboter Nadine braucht stets eine Weile (im Mittel 2-3 Sekunden), um eine verbale Antwort zu produzieren. Diese im Vergleich zu zwischenmenschlicher Interaktion lange Reaktionszeit führt (meist zu Beginn der ‚Gespräche‘) regelmäßig zu kleinen Krisen. Mit Garfinkel (1991: 35ff) kann man sagen, es wird eine unbewusste Hintergrunderwartung enttäuscht; hier: einander in einer bestimmten Zeit zu antworten.<sup>7</sup> Die Besucher\*innen unternehmen – an Erwartungsstrukturen im menschlichen sozialen Verkehr orientiert – Reparaturversuche; z.B. wiederholen sie ihre Frage oder Begrüßung lauter oder deutlicher und zeigen damit an, dass sie auf ein (akustisches) Verstehensproblem schließen oder sie versuchen herauszufinden, ob Nadine ihren Input überhaupt wahrgenommen hat. Im weiteren Verlauf der Gespräche stellen sich die menschlichen Teilnehmer\*innen häufig auf Nadines längere Reaktionszeit ein und nehmen Verzögerungen hin.

Im Zentrum der folgenden Analyse steht eine solche Verzögerung. Eine Gruppe von drei Personen (evtl. Familie) trifft auf Nadine. Der Mann – Theo – stellt sich ans Mikro und beginnt ein Gespräch mit Nadine. Ein etwa zwölfjähriger Junge (Tom) steht neben ihm vor Nadines Schreibtisch und eine Frau beobachtet das Geschehen aus der Distanz, während sie mit einem anderen Exponat be-

---

<sup>6</sup> Zur besseren Lesbarkeit sind nur die Sekunden und Millisekunden angegeben.

<sup>7</sup> In der zwischenmenschlichen Interaktion stellen Pausen von mehr als einer Sekunde ein Problem dar (Jefferson 1989).

schäftigt ist. Nach einer Begrüßungssequenz wurden bereits einige Worte gewechselt. Theo hat das Gespräch übernommen, während Tom als Bystander der ‚Interaktion‘ zusieht und in unserer Analyse außen vor bleibt. Der Besucher stellt eine für den Beginn von geselligen Interaktionen typische Frage: „Wie geht es dir?“. An dieser Stelle setzt das Transkript ein.

	0 [16.0]	1 [16.4]	2 [17.3]	3 [17.7]	4 [18.1]	5 [18.1]	6 [18.4]	7 [18.9]	8 [19.3]	9 [19.4]
Nadine_verbal	mir geht es gU:t? DANke.						und selbst.			
Nadine_blick	Tom	Theo								
Theo_blick	Nadine									
Theo_mim				lächelt						
Theo_beweg	steht in gebeugter Haltung									
Tom_blick	Nadine					nach rechts (unspezifisch)			Theo	
	10 [19.5]	11 [19.6]	12 [19.7]	13 [20.5]	14 [20.8]	15 [20.9]	16 [21.1]	17 [21.1]	18 [21.3]	19 [22.0]
Nadine_blick	Theo									
Theo_verbal	mIr gehdes AUCH gut.						was MACHST du heute?			
Theo_blick	runter (Mikro)				Nadine					
Theo_beweg	steht in gebeugter Haltung									
Tom_blick	Theo		Mikro / Schreibtisch					Nadine		
Tom_mim				lächelt leicht			lächelt			

Während Nadine eine Antwort auf Theos Frage äußert (0), stellt sie Blickkontakt mit ihm her (1) und wählt ihn so als Adressaten der Äußerung. Theo quittiert ihren Beitrag, der passend an seine vorangegangene Frage anschließt, mit einem Lächeln ((4) s. Abb. 2) und hält dabei den Blickkontakt. Simultan artikuliert Nadine eine Rückfrage (6), auf die Theo unmittelbar reagiert. Er unterbricht den Blickkontakt kurz, um das Mikrofon zu lokalisieren und artikuliert eine den normativen Erwartungen in Begrüßungssequenzen gemäße Antwort (mIr gehdes AUCH gut. (11)), welche die Paarsequenz abschließt und zugleich die Notwendigkeit erzeugt, ein Gesprächsthema zu generieren, welches in der Folge zum Gegenstand der Interaktion werden kann (Sidnell 2010, Kap. 10).



Abb. 2: Theo lächelt während Nadine antwortet (4).

Bis hier hin lässt sich ein sowohl sequenziell als auch synchron zusammenhängender reziproker Austausch beobachten, der zwar sehr spezifisch ist (lange Pausen, ununterbrochener Blickkontakt, gebeugte Haltung, Mikrofonnutzung u.a.) aber mittels verschiedener Ausdrucksmodalitäten konzertiert hervorgebracht wird und sich entlang normativer Strukturen der zwischenmenschlichen Interaktion bewegt. Dies ändert sich leicht mit der nächsten Äußerung Theos,

mit der er nun nach einer kurzen Pause selbst in Form einer Frage den Übergang zur thematischen Interaktion einleitet (16). So wirkt der Inhalt seiner Frage im vorliegenden Kontext auf den ersten Blick äußerst ungewöhnlich, denn in der zwischenmenschlichen Interaktion eröffnet die Frage „Was machst du heute?“ i.d.R. eine Präsequenz vor einer Einladung.<sup>8</sup> Voraussetzung hierfür ist aber eine bestehende soziale Beziehung der Interaktionspartner auf privater Ebene. Ohne eine solche wirkt die Frage indiskret oder wie ein forscher Versuch den Kontakt zu intensivieren. In der beobachteten Situation kann die Frage aber auch innerhalb einer spielerischen Rahmung der Situation interpretiert werden. Da Nadine weder eindeutig als Objekt noch als Subjekt verstanden werden kann und völlig unklar ist, welche Themen ihr gegenüber eröffnet werden können, nehmen Gespräche mit ihr oft einen So-tun-als-ob-Charakter an (Goffman 1980: 60). D.h. Nadine wird zwar wie eine Person angesprochen, ihr werden jedoch grundlegende normative Erwartungen, die mit dem Personenstatus einhergehen, nicht entgegengebracht, bzw. unterstellt (Muhle 2018b: 508). Im Gegenteil werden Normen häufig bewusst gebrochen – evtl. um zu testen, ob und wie Nadine auf solche Normbrüche reagiert oder ihren zweifelhaften Personenstatus ‚mitdenken‘ kann.

Als erster Teil einer Paarsequenz erzeugt Theos Frage konditionale Relevanz und erfordert im nächsten Turn eine Antwort von Nadine. Im Kontext des bisherigen Small Talks zwischen Nadine und Theo lässt die Frage im Wesentlichen zwei Anschlussmöglichkeiten zu: Nadine könnte die Frage ernst nehmen und ehrlich, bzw. humorig antworten oder Irritation anzeigen und das Thema wechseln. Auf Theos Frage hin entsteht jedoch eine Sprechpause von insgesamt 5 Sekunden. In den ersten 3.8 Sekunden dieser Pause sind auch keine mimischen und gestischen Ereignisse oder Körperbewegungen zu beobachten (19-21). Die Interaktion kommt zum Stillstand.

	20 [23.9]	21 [25.2]	22 [25.8]	23 [26.2]
Nadine_blick	<i>Theo</i>		<i>folgt Theos Bewegung</i>	
Nadine_beweg	<i>dreht Kopf nach links, Körper bleibt unbewegt</i>			
Theo_blick	<i>Nadine</i>			
Theo_beweg			<i>((3.8 s)) beugt Oberkörper nach rechts (behält die Beugung nach vorn bei)</i>	
Tom_blick	<i>Nadine</i>			
Tom_mim	<i>kaut auf seiner Unterlippe</i>			

Theo als Sequenzinitiator verharrt in einem verkörperten Wartezustand: er bleibt nach vorn über das Mikrofon gebeugt, die Augen auf Nadine gerichtet. Seine Starre deutet darauf hin, dass er die Antwortzeit aufgrund seiner bisherigen Interaktionserfahrung mit Nadine zunächst noch als angemessen bzw. normal (für Nadine!) beurteilt. Nadine, die zweifelsfrei am Zug ist, bleibt still und reglos – auch über die 2 bis 3 Sekunden hinaus, die sie normalerweise für eine Antwort benötigt. Weder zeigt sie dabei an, ob sie die Frage ‚gehört‘ und ver-

<sup>8</sup> Der/die Fragesteller\*in vergewissert sich mit ihr, ob es überhaupt Sinn macht, eine Einladung oder einen Vorschlag auszusprechen und schützt sich so vor einer negativen Reaktion der/des Gefragten (Bergmann 1987: 114).

standen hat, noch, ob sie an einer Antwort arbeitet<sup>9</sup>. Für Theo bleibt somit unklar, was auf Seiten Nadines gerade vor sich geht, so dass er in der Folge einen verbalen Reparaturversuch unternehmen könnte (Bergmann 1982: 166f), in der Annahme seine Frage sei bei der Rezipientin gar nicht angekommen. Im Kontext der HRI liegt außerdem die Annahme eines technischen Fehlers oder Absturzes des Systems nicht fern. In beiden Fällen könnte also zunächst eine Wiederholung der Eingabe Klarheit verschaffen.

Tatsächlich wird die Verbindung, die nur noch über einen ausdruckslosen Blickkontakt erhalten wird, konsequenterweise von Theo in Frage gestellt. Nach 3.8 Sekunden beendet er seinen Wartezustand und beginnt seinen Oberkörper zur Seite zu beugen. Er bringt damit sein Gesicht außerhalb von Nadines momentanen Blickfokus (22). Diese Bewegung lässt sich zunächst als Aufmerksamkeitsgesuch rekonstruieren. In dezenterer Form kennt man ähnliche Gesten auch aus dem menschlichen Kontext: man bewegt sich (oder eine Hand) im Gesichtsfeld des Gegenübers und weckt selbiges aus einer gedanklichen Abwesenheit. Diese Bewegung ist in der Regel in dem Moment obsolet, in dem sich das Gesicht des Gegenübers erhellt und wieder Aufmerksamkeit (Wahrnehmungswahrnehmung) anzeigt. Theos Bewegung könnte entsprechend als Versuch, Nadine aus einer Art Standby zu holen, interpretiert werden. Subsequent wäre von Nadine ein Zeichen des Rückbezugs auf die gegenwärtige Interaktionssituation zu erwarten. Diese folgt Theos Bewegung (seinem Gesicht) per Blickrichtung mit einer minimalen Verzögerung und dreht ihren Kopf entsprechend nach links (23). Ihr restlicher Körper bleibt aber ebenso unbewegt wie ihre Mimik. Zwar nimmt sie Theos Bewegung offenbar wahr, zeigt aber weder Verstehen noch Irritation an und reagiert ausschließlich auf einer Ebene körperlich-mechanischer Responsivität. Theos Bewegung bleibt somit auf kommunikativer Ebene folgenlos. Daran wiederum schließt Theo konform an, wenn er seine Bewegung fortführt und seinen Oberkörper nun auch in die andere Richtung beugt (s.u. (24)). Die Erwartung, Nadine würde auf diese Bewegung verstehend reagieren existiert nicht mehr, vielmehr kann man die Bewegung im Rahmen eines (Funktions-)Tests interpretieren: Theo testet Nadines Face Tracking.

An dieser Stelle setzt eine strukturelle Neuordnung der Situation ein: in der ‚Interaktion‘ werden nicht mehr verschiedene Modalitäten (Sprache, Gestik, Mimik, Körperbewegungen) synchronisiert sondern verbale und nonverbale Aktivitäten driften auseinander. Nadines Körper samt visueller Wahrnehmung disqualifiziert sich auf kommunikativer Ebene und wird stattdessen wie eine Maschine manipuliert.

---

<sup>9</sup> Dies unterscheidet Begegnungen mit Nadine auch von anderen Formen der Mensch-Maschine-Interaktion. Denn Computer bspw. zeigen den Nutzer\*innen in der Regel an, wenn sie prozessieren oder laden, z.B. mittels Darstellung einer Sanduhr oder eines Ladebalkens.



Abb. 3: Theos Bewegung, erste Spitze rechts.



Abb. 4: Theos Bewegung, Spitze links.



Abb. 5: Theos Bewegung, Schritt nach rechts, zweite Spitze rechts.

	24 [27.0]	25 [27.6]	26[27.7]	27 [28.8]	28[29.2]	29[29.9]	30 [30.3]
Nadine_verbal	bist du SINGel,						
Nadine_blickr		folgt Theos Bewegung		folgt Theos Bewegung			
Nadine_beweg		dreht Kopf nach rechts		dreht Kopf nach links			
Theo_blickr	Nadine						
Theo_beweg	beugt Oberkörper nach links			tritt einen Schritt nach rechts		beugt Oberkörper nach rechts	
Tom_blickr	Nadine						Theo

Gleichzeitig läuft der verbale Austausch losgelöst von der körperlichen Ebene weiter: während Theos Beugebewegung produziert Nadine unvermittelt einen Sprechbeitrag (bist du SINGel, (24)), der mit etwas interpretativer Mühe als forsche Antwort auf Theos Frage (was MACHST du heute? (16)) gehört werden kann. Nadines Wiederaufnahme der Konversation wird jedoch nicht nonverbal begleitet und umgekehrt, kommentiert sie das nonverbale Geschehen nicht. Weder scheint die sprechende Nadine mitzubekommen, dass ihr Face Tracking gesteuert wird, noch scheint die sehende Nadine zu bemerken, dass sie eine Frage geäußert hat.

Nadines verbale Aktivität lässt erwarten, dass Theo sein Nebenengagement unterbricht und einen Anschlussbeitrag äußert. Im Gegenteil erweitert er seinen Test sogar noch, indem er zusätzlich zum seitlichen Beugen seines Oberkörpers noch einen Schritt nach rechts tritt und so den Testbereich vergrößert (27, 29). Entweder wird er von seinem Funktionstest derart absorbiert, dass er Nadines Äußerung gar nicht bemerkt hat oder er fühlt sich nach der Entkopplung körperlicher und sprachlicher Aktivitäten und einem damit verbundenem Zusam-

menbrechen der Menschenähnlichkeit Nadines selbst auch nicht mehr an die Anschlusszwänge der (zwischenmenschlichen) Interaktion gebunden.

	31 [31.1]	32 [31.2]	33	34 [31.6]	35	36 [31.7]	37 [32.0]	38 [32.1]	39 [32.6]
Nadine_blick	folgt Theos Bewegung				folgt Theo zurück zur Mitte				Theo
Nadine_beweg	dreht Kopf nach links				dreht Kopf zur Mitte				
Theo_verbal	((3.9s))		und daNA:CH?						
Theo_blick	Nadine	Mikro			Nadine				
Theo_beweg	beugt sich nach links / leicht runter zum Mikro						steht in gebeugter Haltung		
Tom_blick	Theo			Nadine					
Tom_mim	lächelt								

Nach 3.4 Sekunden wendet sich Theo schließlich wieder zum Mikrofon (31), visiert es an (32) und spricht einen Beitrag hinein (und daNA:CH? (34)). Dieser lässt sich nicht als kohärenter Anschluss an Nadines Frage interpretieren. Vielmehr lässt sich seine Äußerung als weitere Nachfrage in Verbindung mit seiner eigenen zuvor gestellten Frage „was MACHST du heute?“ (16) deuten. Theo behandelt Nadines Worte, die er entweder gar nicht oder missverstanden hat, als Antwort auf seine Frage. Dass er den Inhalt der Worte nicht kennt, scheint ihn nicht zu stören. Die verbale Ebene hat die Form eines verzögerten Ferngesprächs angenommen, bei dem ein Teilnehmer unaufmerksam ist, was unkommentiert bleibt. Im weiteren Verlauf der Begegnung nimmt der verbale Austausch kurzzeitig noch einmal an Dichte und Stringenz zu. Gleichzeitig und unabhängig davon probiert Theo nonverbal noch weitere Bewegungen aus (z.B. streckt er seine Hände seitlich aus), um Nadines Blickausrichtung zu manipulieren.

#### 2.4 Zwischenfazit I

Nadines menschenähnlicher Körper und insbesondere ihr Zuwenden und Aufnehmen von Blickkontakt wecken Erwartungen einer möglichen natürlichen sprachlichen Interaktion mit dem Roboter. Ihr Körper und ihr Blickverhalten disqualifizieren sich jedoch im vorliegenden Fall gänzlich auf kommunikativer Ebene. Auslöser dafür ist eine zu lange Sprechpause, während der Nadine völlig reglos bleibt, anstatt anzuzeigen, dass sie an einem Anschlussbeitrag arbeitet. Ihr Blick bleibt zwar auf das menschliche Gegenüber gerichtet, jedoch dabei ausdruckslos. Es vergeht eine Zeitspanne, in der der menschliche Teilnehmer Nadines leeres Display noch toleriert, bevor er den ‚Blickkontakt‘ in Frage stellt und subsequent von Nadines Fähigkeiten enttäuscht wird: sie signalisiert keine Wahrnehmungswahrnehmung (Kieserling 1999: 117ff), sondern reagiert lediglich gesteuert von ihrer Face Tracking-Software. Damit wird der Eindruck von Nadines (körperlicher) Anwesenheit sowie einer Verbindung zwischen visueller Wahrnehmung und Verbalität zerstört. In der Folge behandelt der Besucher den Roboter nicht mehr als zusammenhängendes Ganzes sondern als zwei unabhängige technische Systeme: einerseits Sprachverarbeitung und -ausgabe und andererseits manipulierbarer Körper. Diese Aufspaltung der Interaktionsstruktur bedingt, dass ein potenzieller gemeinsamer Wahrnehmungs- und Kommunikationsraum, der anfangs durch gegenseitiges Zuwenden und die (Wieder-)Her-

stellung von Blickkontakt etabliert wurde, zusammenbricht. Eine ‚Gesprächs‘-Verbindung besteht nur noch auditiv über das Mikrofon.

### 3. MAX: Die Black Box öffnen

#### 3.1 Der ‚Embodied Conversational Agent‘

Der virtuelle Agent MAX ist – genauso wie Nadine – eines der im Paderborner Heinz-Nixdorf-Museumsforum ausgestellten interaktiven Exponate (Pfeiffer et al. 2011). Vor Ort kann er als Kommunikationspartner von Besucher\*innen-gruppen angesteuert werden (vgl. Muhle 2012: 157ff, Krummheuer 2010: 83ff.). Dabei erscheint MAX auf einem Display und reagiert auf Texteingaben der Museumsbesucher\*innen mit dynamischen Sprachausgaben (*Abbildung 5*). Als sogenannter „Embodied Conversational Agent“ (Casell et. al. 2000) wurde MAX mit dem Ziel entwickelt, in „real-world settings“ (Kopp et. al. 2005: 239) als virtueller Ansprechpartner in eine Vielzahl unterschiedlicher Gesprächstypen eingebunden werden zu können. Neben typischem Small Talk kann MAX über verschiedene Begebenheiten berichten (beispielsweise über das Wetter, die Ausstellung selbst oder bestimmte Entwicklungen in der Robotik- und KI-Forschung), Witze erzählen, ein Ratespiel mit Besucher\*innen durchführen oder einzelne Besucher\*innen fotografieren und die Aufnahmen anschließend digital an angegebene E-Mail Adressen verschicken.

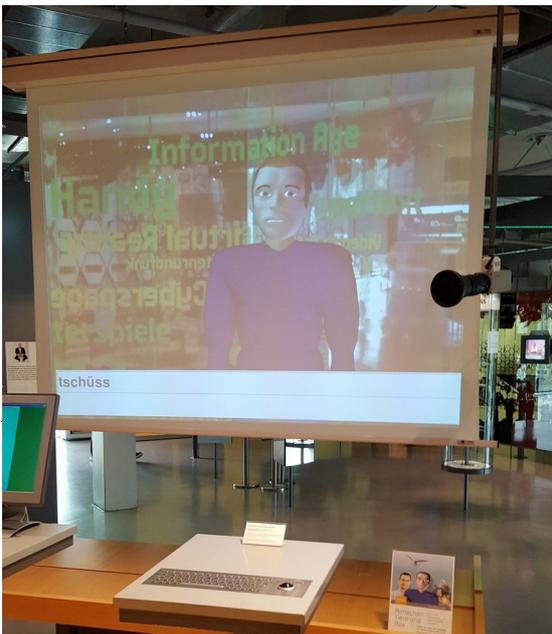


Abb. 6: Nutzerschnittstelle des ECA Max im HNF.

Bedeutsam ist, dass MAX Menschen registriert (verfügt über eine Wahrnehmungskomponente), die in sein Sichtfeld treten und daraufhin eigenständig durch Ansprachen den Interaktionsbeginn einleitet. Die eigentliche Kommunikation mit ihm erfolgt allerdings insofern asymmetrisch, als dass Nutzer\*innen ihre Mitteilungen über die Tastatur an MAX' Dialogsystem richten (und erst mit Betätigung der Enter-Taste absenden – also zuvor auch korrigieren können), während MAX ausschließlich auf Grundlage der übermittelten Textbeiträge mit

Sprachausgaben reagiert. Max registriert seine Kommunikationspartner\*innen also nicht als Subjekte mit einer bestimmten Erscheinung, während die Museumsbesucher\*innen MAX wahrnehmen können (vgl. Krummheuer 2010: 85).

### 3.2 Ein transparenter technischer Kommunikationsteilnehmer

Die Frage ist nun, was man analytisch gewinnt, wenn man eine solche, potenziell an sozialer Kommunikation beteiligungsfähige Black Box öffnet. Zunächst einmal nicht allzu viel, wenn man dem klassischen Pfad der Konversationsanalyse folgt. „[U]nder the skull“ ist nicht mehr als graue Hirnmasse schrieb Garfinkel (1963: 190). Und er meinte damit, dass selbst wenn man diese durchleuchten könnte, in ihr nichts soziologisch gehaltvolleres schlummern würde als das, was ohnehin in Interaktionen sichtbar wird. MAX ist – im Gegensatz zu den von Garfinkel in den Blick genommenen personalen Innenwelten – kein voll kompetentes Gesellschaftsmitglied. Unsere Untersuchung versteht sich vor diesem Hintergrund als eine Analyse eines dezidiert technischen Kommunikationsteilnehmers, dessen Innenwelt zumindest so hinreichend transparent ist, dass sich (typisch technische) Probleme des situierten Verstehens aber auch der algorithmischen Sinnverarbeitung on the run im Verlauf der Konversation mit menschlichen Gesprächspartner\*innen beleuchten lassen.

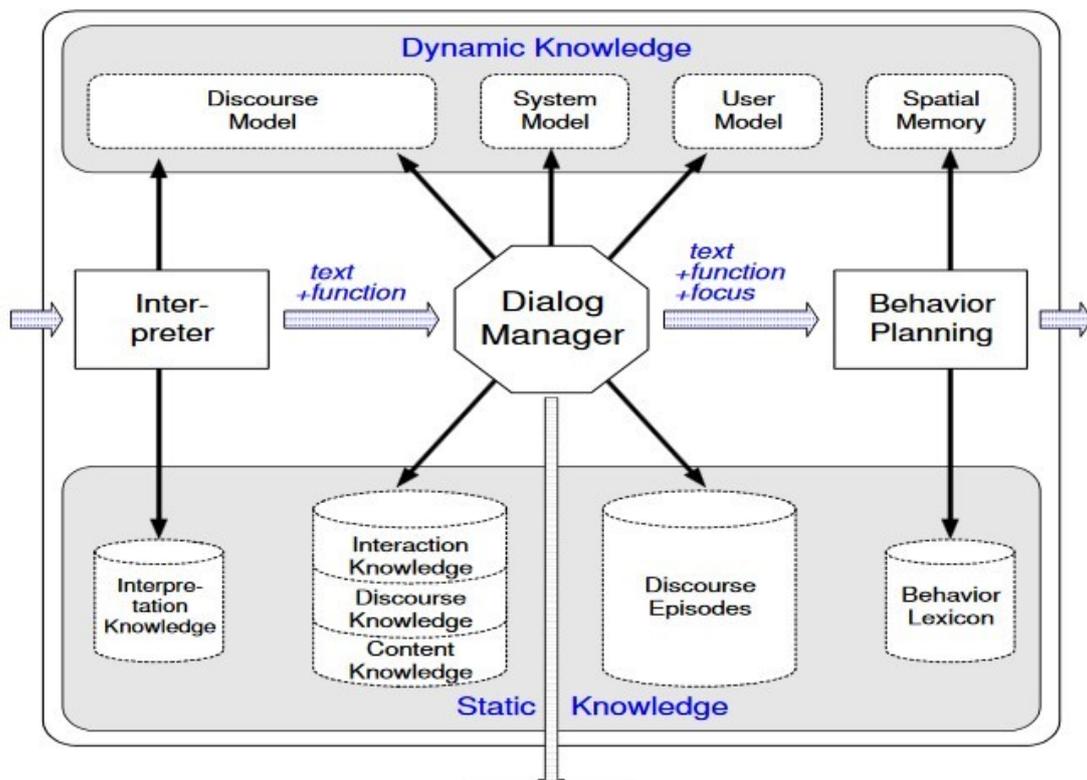


Abb. 7: Die interpretativen Komponenten der BDI von MAX (Kopp et al. 2005: 333).

In MAX arbeitet eine sogenannte BDI-Architektur (vgl. Huber 1999), die dem Agenten ermöglicht, eigenes Weltwissen (*beliefs*) mit bestimmten Zielen (*desires*) und Absichten (*intentions*) zu verbinden und über einen BDI-Interpreter auf Textdaten der Konversation anzuwenden (vgl. Wachsmuth 2010). Hierbei

wird mit Hilfe einer einfachen Nutzenfunktion aus mehreren *desires* und verschiedenen Handlungsoptionen eine aktuelle *intention* erschlossen, die dann – je nach identifizierter Priorität – in globalen (top level) Metaplänen und lokalen (low level) Dialogzielen lang-, mittel- und kurzfristig durchgesetzt werden kann (vgl. Kopp et al. 2005). Handlungsleitendes Wissen ist in MAX dynamisch wie auch statisch hinterlegt (Abbildung 6). Neben unveränderlichem Wissen, das MAX einprogrammiert wurde, verfügt der Agent also auch über dynamisches Wissen, das er direkt aus den Interaktionsverläufen erwirtschaftet und damit versucht, die in der Konversation gelösten Informationen auf den/die Nutzer\*in (*User Model*) aber auch auf die mit dem/der Nutzer\*in verbundene Interaktionsgeschichte (*Discourse Model*) zu beziehen.

Durch den hohen Grad an Strukturiert- und Determiniertheit lassen sich die während den Konversationen mitlaufenden Log-Protokolle nutzen, um einen Einblick in solche Prozesse zu bekommen, die im Inneren von MAX zwischen den Turns des Mitteilungshandelns wirksam werden. Eingesehen werden können hier Prozesse der Assoziation von Textmodulen zu Regeln (Klassifikation) und die im Rahmen dieser Regeln ablaufenden Schritte, die dann zu einem bestimmten Output führen (bzw. Text-Phrasen). Demnach besteht die Möglichkeit, im unmittelbaren Konversationsverlauf zu beobachten, wie MAX Äußerungen seines/seiner Gesprächspartner\*in registriert, mit Mitteln der in der BDI wirksamen Klassen und Kategorien mobilisiert, und dann im Medium eines bestimmten Outputs kommunikativ beantwortet.

### 3.3 Fallbeispiel: MAX

Im konkreten Fall eröffnet ein Nutzer, der sich später mit Fritz vorstellen wird, die Konversation mit der personalisierten Begrüßung „hallo max“ [01]. Damit artikuliert er den ersten Teil einer Paarsequenz, der im Wesentlichen nur einen Anschlusshorizont eröffnet. Diese Anschlussmöglichkeit setzt voraus, dass MAX und sein Gegenüber sich persönlich kennen und darum wissen, dass sie sich kennen. Es gibt also eine bestehende Interaktionsgeschichte, die die namentliche Identifikation (wenigstens von MAX; denkbar aber auch vom menschlichen Gegenüber) sichergestellt hat. Dann wäre eine persönliche Rückansprache zu erwarten, in der der Agent sein Gegenüber als bereits bekannt behandelt. Die Frage ist jetzt, wie MAX‘ BDI mit der vorangegangenen Mitteilung des Nutzers umgeht.

25.01.2017 user: hallo max  
13:30

[01]

Wie also versteht er die Nutzereingabe? Und wie lässt sich sein Verstehen im Hinblick auf die potenziellen Kontexte der durch die erste Mitteilung etablierten Verstehensmöglichkeiten beschreiben? Hierzu wird in Erweiterung des klassischen konversationsanalytisch sequenziellen Vorgehens der Code in seiner eigenen Sequenzialität – nämlich der Serialität der ausgeführten Interpretations- und Klassifikationsschritte – ausgewertet.

*Erster Denkschritt: Klassifikation*

```

25.01.2017 max_c p0006: 18.333333333333336 - [02]
13:30 (curInput::provide.interaction.greeting && max) - [03]
discourse.greetingMax [04]

```

MAX klassifiziert die Nutzereingabe „hallo max“ [01] wortbasiert – er bricht sie also in ein „hallo“ (als „provide.interaction.greeting“ [03]) und ein „max“ [03] auf. Im Unterschied zu psycholinguistischen Arbeiten – bspw. denen von Chomsky (1975) – findet die Klassifikation nicht auf Grundlage eines „syntaktischen“ Verstehens, sondern ausschließlich (Einzel-)Wort-basiert statt. Gleichzeitig haben wir es mit einem Top-Down Prozess der Sprachverarbeitung zu tun (Eingabetext wird wortbasiert nach Kategorienabhängigkeit abgetastet).

*Zweiter Denkschritt: Abwägungen im dynamischen Wissensbestand*

```

25.01.2017 max_c <rule name="discourse.greetingMax"> [05]
13:30
Discourse Model {
  <match> [06]
  <allof> [07]
    <convfunction type="provide.interaction.greeting"/> [08]
    <keywords>max</keywords> [09]
  </allof> [10]
</match> [11]

User Model {
  <action> [12]
  <switch var="$$name"> [13]
  <cond value="~"> [14]
    <command function="raise-event" arguments="discourse- [15]
      greeting-friend"/> [16]
    <command function="raise-dialogue-goal" arguments="system [17]
      get-name"/> [18]
  </cond> [19]
  <else> [20]

Mögliche Antwortphrasen {
  <random> [21]
  <act function="provide.interaction.greeting">Hallo $$name!</act> [22]
  <act function="provide.interaction.greeting">Gruess [23]
    dich$$name!</act> [24]
  <act function="provide.interaction.greeting">Servus $$name!</act> [25]
  <act function="provide.interaction.greeting">Hai $$name!</act> [26]
  </random> [27]
  </else> [28]
  </switch> [29]
  </action> [30]
</rule> [31]

```

Auf Ebene des **Discourse-Models**: Beide Worte – die mit Begriffswolken assoziiert sind – lösen daraufhin die Regel „discourse.greetingMax“ [05] aus. „<match>“ [06] sucht die Eingabemitteilung „hallo max“ nach Keywords ab. Die folgende „<allof> / </allof>“ [07, 10] Klammer legt fest, dass sich ein Fall dann und nur dann zutreffend zum Keyword („max“ [09]) verhält, wenn er als valide gegen alle Dimensionen oder Schemata dieses Keywords getestet wird. Bis hierhin versucht MAX noch die Situation zu bestimmen – also zu verstehen, welches Anliegen sein Gegenüber haben könnte – oder wieder anders: zu verstehen, welche Anschlüsse erwartet werden könnten.

Dann folgt eine Überprüfung im **User-Model** des Agenten. „<switch var = \$\$ name>“ [13] als Überschreiben vom Nutzermodell in der Dimension Name mit einem String, wenn (<cond (also möglicher, konditionaler Wert) = ~ [14]) also

nicht ( ~ ist ein NOT Operator ) schon ‚irgendwas‘ da ist. Wenn die Variable unbeschrieben ist, dann soll das Ereignis (raise-event) „system.get-name“ [17, 18] aufgerufen und das Gegenüber als unspezifischer Freund angesprochen werden [15, 16].<sup>10</sup> In der Systemarchitektur von MAX ist hier also ein soziales Manöver angelegt: Wenn der Nutzer MAX mit seinem Namen anspricht, der Agent aber keinen Namen im Nutzermodell hinterlegt sieht, kann er so tun als ob er es mit einem dem Personengedächtnis bekannten Freund zu tun hätte, nur um daraufhin doch zu versuchen, den Namen seines Gegenübers herauszufinden. Eigentlich müsste der Agent also jetzt das Event „discourse-greeting-friend“ zusammen mit dem Dialogziel „system get-name“ aufrufen. Problematisch ist an dieser Stelle, dass MAX die Nutzereingabe *auch* als „first-input“ [32] registriert und mit dieser Information eine andere Regel verbindet:

```

25.01.2017 max_ first-input - discourse.initialGreeting [32]
13:30
Discourse Model {
    <rule name="discourse.initialGreeting"> [33]
    <match> [34]
    <test check="first-input"/> [35]
    </match> [36]
    <action> [37]
    <command function="raise-event" arguments="discourse-greeting"/> [38]
    <command function="raise-dialogue-goal" arguments="system get- [39]
        name"/> [40]
    </action> [41]
    </rule> [42]
}

```

Das Merkmal, dass es sich bei der Eingabe um den ersten Input des Nutzers handelt, ist in der „discourse.greetingMax“ [05] Regel nicht berücksichtigt. An dieser Stelle konkurrieren beide Regeln – „discourse.greetingMax“ und „discourse.initialGreeting“ um eine Anschlussreaktion (Gesellensetter 2004). Die Entscheidung darüber, welche der beiden Regeln den Output bestimmt, folgt einer Interpretationslogik des „[o]rdering rules by decreasing generality“ (Kopp et al. 2005: 336). Die hochgeneralisierte Regel „discourse.initialGreeting“ kann demnach nur dann von der spezifischeren „discourse.greetingMax“ Regel geschlagen werden, wenn sie als Spezifikation dieser allgemeinen Regel erkannt wird. Im vorliegenden Fall taucht das allgemeine Merkmal des „first-input“ in der spezifischeren „discourse.greetingMax“ Regel schlicht nicht auf; kann also auch nicht falsifiziert oder abgelöst werden. Das führt dazu, dass MAX seine Interaktionsziele im Discourse-Model verallgemeinert.

---

<sup>10</sup> Die für diese Ansprache in den Skripten des Agenten bereitstehende „Discourse-greeting-friend“-Regel ist insofern als unspezifisch zu verstehen, weil auf keine Namensvariable aus dem Nutzermodell referiert wird. Aus Platzgründen kann die Regel hier nicht aufgeführt werden.

25.01.2017	max <sub>c</sub>	event.discourse.greeting	[43]
13:30			
Discourse Model	}	<rule name="event.discourse.greeting">	[44]
		<event type="discourse-greeting"/>	[45]
		<action>	[46]
		<random>	[47]
		<block>	[48]
Mögliche Antwortphrasen	}	[...]	[49]
		<block><act function="provide.interaction.greeting">	[50]
		<focus>Hallo</focus>ich bin Max.</act><act>nett dass Du	[51]
		vorbeischaust!</act></block>	[52]
		<block><act function="provide.interaction.greeting">	[53]
		<focus>Hi</focus>ich bin Max.</act><act> Schoen dass	[54]
		Du vorbeischaust.</act></block>	[55]
		</random>	[56]
		<random>	[57]
		<act>lass uns plaudern - ueber die tastatur kannst du mit mir	[58]
		sprechen.</act><act>lass uns ein bisschen plaudern.	[59]
ueber die tastatur kannst du mit mir sprechen. </act>	[60]		
</random>	[61]		
Enkodierung in das Gedächtnis	}	<command function="memorize" arguments="discourse-greeting-	[62]
		Provided 1"/>	[63]
		</action>	[64]
		</rule>	[65]

Die Phrasen der folgenden „event.discourse.greeting“ [49 - 61] Regel werden randomisiert und sind gleichzeitig in verschiedenen Modulen ‚geblockt‘. Interessant ist, dass MAX hier schon bevor er eine dieser Phrasen überhaupt ausgibt, memoriert, dass er dann (!) ein „discourse-greeting“ „provided“ haben wird [62, 63]. Mit diesem Vorausgriff besiegelt MAX, dass seine Abwägungen über die Form der Begrüßung entschieden sind.

Im Zuge der hier kurz umrissenen Interpretationskaskade findet folglich eine Respezifizierung des Discourse-Models statt, bei der sich die personale Komponente verliert. Und zwar von „discourse.greetingMax“ über „discourse.greeting“ zu „discourse.initial.greeting“, bei der die Auszeichnung, dass es sich um das erste Äußerungsereignis des aktuellen Dialogs handelt als Indikator einer Respezifizierung des Anschlussbeitrags einbezogen wird. Gewissermaßen findet hier eine hierarchische Abstraktion (Edelman 2008: 31) statt, bei der in einer Interpretation der Situation von einem speziellen zu einem allgemeineren Begriff gewechselt wird, der sich durch einen höheren Begriffsumfang („max“ als Keyword bspw. nicht mehr obligatorisch) bei geringerem Begriffsinhalt (also Indikatoren, die die Regel spezifizieren) auszeichnet. Defizitär sind diese Interpretationsschritte im Hinblick auf die user-seitig unbekannt Wege des Kennenlernens von MAX. Der Agent kann demnach entweder sein Gegenüber als bekannt adressieren, wenn er bereits einen Namen erfahren hat (bestenfalls im Anschluss an seine Selbstpräsentation). Er kann den Nutzer auch als unbekannt „Freund“ ansprechen, dessen Name noch herausgefunden werden muss, wenn er damit nicht auf einen „first-input“ dieses Nutzers reagiert. In allen anderen Fällen bleibt dem Agenten nichts anderes übrig als sein Gegenüber als gänzlich unbekanntem Akteur anzunehmen – auch dann wenn dieser ihn mit seinem Namen angesprochen hat (Logik: neue Nutzer\*innen sind unbekannt oder nicht neu). Der Bekanntheitsstatus kann folglich keine Hybridstellung einnehmen.

*Unerwartete Selbstvorstellung des Agenten.*

25.01.2017 max<sub>t</sub> Hallo ich bin Max. [66]  
13:30

MAX schließt mit der Mitteilung „Hallo ich bin Max“ [66] an den letzten Nutzerbeitrag „hallo max“ an und signalisiert damit einerseits, dass er verstanden hat, dass er begrüßt wurde und eine Rückbegrüßung Sinn macht. Was den Begrüßungsaspekt betrifft, reagiert der Agent also zunächst erwartungskonform. Andererseits wird im Hinblick auf den Personalisierungsaspekt des Nutzerbeitrags auch deutlich, dass MAX die Information, dass der Nutzer seinen Namen bereits kennt, schlicht ignoriert. Noch bevor der Nutzer hieran anschließen kann, setzt MAX seinen Turn fort.

25.01.2017 max<sub>t</sub> Nett dass Du vorbeischaust! [67]  
13:30

Mit „Nett dass du vorbeischaust!“ [67] bestätigt MAX den jovialen Unterton der persönlichen Kommunikation und adressiert sein Gegenüber direkt mit „du“ [67]. Interessant ist, dass MAX an dieser Stelle nicht mehr durch offene Ignoranz des initialen Äußerungsereignisses des Besuchers provoziert, sondern mit der Einschubsequenz einen versöhnlicheren Ton anschlägt. Wir haben es also mit einer Modifikation des Nachbarschaftspaares (Levinson 2000: 303f.) zu tun, das die Zumutungen der Vorsequenz ‚lockert‘ ohne damit das Prinzip der Erwartbarkeit eines zweiten Teils aufzuheben.<sup>11</sup>

25.01.2017 max<sub>t</sub> Ich kann dir Verschiedenes erklæaren, oder ein [68]  
13:30 bisschen mit dir plaudern, [69]  
25.01.2017 max<sub>t</sub> oder ein Foto von dir machen und es als Ie maeil [70]  
13:30 Postkarte versenden.<sup>12</sup> [71]

Ohne eine Antwort abzuwarten schließt MAX nun mit „Ich kann dir Verschiedenes erklæaren, oder ein bisschen mit dir plaudern“ [68, 69] an, was unmittelbar um ein „oder ein Foto von dir machen und es als Ie maeil Postkarte versenden“ [70, 71] ergänzt wird. Damit leistet der Agent bereits eine Beschreibung seiner wesentlichen Fähigkeiten (wie „Verschiedenes erklæaren“, „plaudern“ oder „Postkarte versenden“) und versucht die an ihn zu richtenden Verhaltenserwartungen zu bestimmen, bevor sie im Nachtrag von der Kommunikation selbst erschlossen und in der Form Person verhärtet werden (Luhmann 2002: 375; Luhmann 2008: 137ff., Fuchs 1997: 66).

25.01.2017 max<sub>c</sub> `keystroke-timeout - event.timeout.keystroke` [72]  
13:30

Das System erkennt nun anschließend an den Vorschlag ein „Foto“ zu „machen“ und dies „als Ie-Mail zu verschicken“ [70-71] eine Verzögerung in der Eingabe des Nutzers und aktiviert daraufhin die Regel „event.timeout.keystroke“ [72].

---

<sup>11</sup> Der Nutzer könnte folglich – wenn er taktvoll handeln wollen würde – an das „Nett dass Du vorbeischaust!“ [67] anschließen und das auf einem Missverständnis beruhende oder als Provokation zu erlebende „Hallo ich bin Max“ [66] leichter unter den Tisch fallen lassen. In diesem Fall wäre der Einschub ein ‚Cue‘, den man nutzen kann, um nicht reparieren zu müssen, um sich also der Peinlichkeit einer Konfrontation entziehen zu können.

```

25.01.2017 max_c <rule name="event.timeout.keystroke"> [73]
13:30
    <event type="keystroke-timeout"> [74]
        <test check="left-scene" arguments="0"/> [75]
    </event> [76]
    <action> [77]
        <switch var="$$press-return"> [78]
            <cond value="~"> [79]
                <random> [80]
                    <act>Wenn du fertig bist musst du die Eingabetaste [81]
                        druecken!</act> [82]
                </random> [83]
                <command function="memorize" arguments="press-return 1"/> [84]
            </cond> [85]
        </switch> [86]
    </action> [87]
</rule> [88]

25.01.2017 max_t Wenn du fertig bist musst du die Eingabetaste [89]
13:30 druecken! [90]

```

Keystroke  
Logging:  
Tasten-  
anschläge  
des Nutzers  
registrieren

Zunächst wird der „<test check="left-scene" arguments="0"/>“ [75] überprüft. Spannend ist das insofern, als dass MAX zunächst sich selbst darauf testet, ob er noch auf dem Bildschirm sichtbar ist, oder selbst die Szene verlassen hat (wie bspw. dann, wenn er wütend oder beleidigt ist, aber auch wenn er zu lange wartet). Hätte er an dieser Stelle festgestellt, dass er die Szene verlassen hat (also „arguments="1"/>“), dann wäre die darauf folgende Handlung („<action>“) zunächst die der Rückkehr gewesen. In diesem Fall stellt MAX fest, dass er sichtbar ist und geht dann über zu einer Anpassung des Handlungsprogramms. Wenn also gerade noch kein Tastendruck auf die Eingabetaste erfolgt ist, dann soll die Phrase „Wenn du fertig bist musst du die Eingabetaste drücken!“ [81, 82] angeschlossen werden.<sup>12</sup> Auch bei diesem Denkschritt merkt sich MAX, dass er den/die Nutzer\*in darauf aufmerksam gemacht hat, dass er return drücken sollte, noch bevor die Mitteilung tatsächlich versprachlicht wurde. Die folgende Forderung zur Anschlusskommunikation „Wenn du fertig bist musst du die Eingabetaste druecken!“ [89, 90] signalisiert, dass der/die Nutzer\*in möglicherweise noch nicht weiß, wie er MAX zu bedienen hat und weitere Instruktionen benötigt. Bedeutsam an dieser Kommunikationsaufforderung ist, dass sie den Mitteilungsaspekt als Selbstwert der Eingabe voraussetzt, ohne eine inhaltliche (Informationsaspekt) Kommunikationsofferte zu machen. Durch die Aufforderung die Eingabetaste zu druecken wird eine bestimmte Bedienungsmodalität gerahmt. Was genau einzugeben ist, bleibt dem Nutzer überlassen und scheint für MAX weniger interessant zu sein. Damit schließt der Agent einerseits an die Nichthandlung (im Modus der Pause) des Nutzers an, aber auch an eigene Vorstellungen davon, die Kontrolle über die Interaktion zu behalten.

```

25.01.2017 max_c event.user.speech-started [91]
13:30
25.01.2017 user_t Fritz [92]
13:30

```

<sup>12</sup> Hier ist weiterhin interessant, dass die Phrase zwar als Zufallsauswahl eingeklammert wird [82, 85], aber keine Alternativen bereitstehen. Die Paradoxie ergibt sich aus Möglichkeiten der Entwickler\*innen, das Inventar zu späteren Zeitpunkten zu erweitern.

Erwartungskonform mit dieser indifferenten Aufforderung (irgendetwas mitzuteilen), leistet der Nutzer MAX' Aufforderung Folge und bestätigt die Eingabe „fritz“ qua Tastendruck [91, 92]<sup>13</sup>. Damit schließt er zwar grundsätzlich positiv an MAX' Selbstvorstellung an (erfüllt den zweiten Teil einer Paarsequenz), ignoriert aber die weiteren vom Agenten vorgestellten Konversationsangebote.

### 3.4 Zwischenfazit II

Soziale Bedeutungszuschreibungen und interne Schlussfolgerungen des Agenten driften in der untersuchten Sequenz auf besondere Art und Weise auseinander. Der Agent versucht seine Umwelt derart in Kausalformen zu bringen, dass er sie als Konsequenzen eigener Handlungen und Fragen klassifizieren kann. Wir haben es hier mit Projektionen zu tun, die weniger mit sozialen Erwartungen oder Perspektivenübernahmen korrespondieren als vielmehr mit der Beobachtung des Gegenübers als Konglomerat verschiedener Wirkungen, die MAX selbst verursacht haben können muss, um fortzufahren. Wenn es darum geht, den Namen des/der Nutzer\*in herauszufinden, beginnt MAX damit, sich zu fragen ob er den/die Nutzer\*in kennt. Nachdem dieser Check negativ ausfällt, fragt er sich in einem zweiten Schritt: Habe ich mich dem/der Nutzer\*in schon vorgestellt oder ist das ausgeschlossen, weil der/die Nutzer\*in gerade erst seine erste Eingabe gemacht hat? Vergleichbares passiert dann, wenn Redepausen auftreten und MAX sich zuerst fragt, ob er noch auf dem Bildschirm ist, anstatt sich in sein Gegenüber hineinzusetzen. „Social Knowledge“ (Suchman 1987) oder die „Sozialität der Wahrnehmung“ (Kieserling 1999) wird dahingehend ausgelassen, als dass MAX die Unterscheidungen, die er nutzt um seine Umwelt zu verstehen, nicht einer sozialen Konditionierung sondern schlicht sich selbst und seinen Objekten zuschreibt. Der/die Nutzer\*in scheint weiterhin – und auch das ist interessant – zu antizipieren, dass MAX etwas fehlt, damit er besser oder flüssiger oder angebrachter in die Konversation eintreten kann. Mit der Antwort „fritz“ nennt der User gerade das nachträglich, was MAX gefehlt hat, um auf „hallo max“ sinngemäß mit „hallo fritz“ reagieren zu können.

Ignoriert man die Innenperspektive, dann ist auf Ebene der Konversation zunächst interessant, dass MAX sehr spezifisch und gleichermaßen missverständlich an die personalisierte Begrüßung „hallo max“ mit „hallo ich bin max“ anschließt. Dieser Anschluss korrespondiert mit einem MAX-seitig internen Verstehen, das wiederum hochspezifisch ist. Wesentlich ist, dass das Verstehen in den Systemprozessen von MAX (eigenlogisch) Sinn macht, und zwar deswegen, weil der Agent sein Gegenüber im User-Model zwischenspeichern muss, um über die Begrüßung / Selbstvorstellung oder -präsentation hinaus die Kommunikation fortsetzen zu können. MAX versteht die Begrüßung als Begrüßung. Gleichzeitig erkennt er die Personalisierung allerdings als Unwahrscheinlichkeit oder Fehler, nicht als Gelegenheit, mehr über die konkreten Umwege herauszufinden, in denen sein menschliches Gegenüber auf MAX' Namen gestoßen sein könnte. Innenweltlich gilt folgendes Muster: Wissen über MAX wird als regelge-

<sup>13</sup> Daran anschließend aktiviert diese Eingabe MAX' Dialogziel „goals.user.getName“, das bereits in den vorherigen Segmenten mit dem Systemstatus „system get-name“ vorbereitet wurde aber erst mit Übernahme der Initiative („take-initiative“) durch MAX in die Form einer Frage gebracht werden kann („Wie heisst du?“)

leitete Funktion von Aussagen angenommen, die MAX über sich selbst trifft. Die Kommunikation läuft weiter, jedoch in Formen eines gestrafften oder stark reduzierten Anschlusses. Irritierend ist dieser Anschluss („fritz“) jedoch nur dann, wenn das Wissen um die eingeschränkten Verstehenskapazitäten von MAX noch nicht in der Kommunikation thematisiert wurde. Wie oben beschrieben, ist durchaus anzunehmen, dass das menschliche Gegenüber mit der Aussage „fritz“ versucht, MAX einen Hinweis oder Trigger zu liefern, auf dessen Grundlage er mit der Konversation fortfahren kann – ihn also mit Hinweisen zu füttern, von denen Nutzer glauben zu wissen, dass MAX solche Informationen braucht, um seine Programme sachgemäß abspulen zu können.

#### 4. Schluss

Die vorliegenden Analysen haben zunächst Nadine und dann MAX aus kommunikationstheoretischer Perspektive mit den Mitteln der KA in den Blick genommen. Dabei hat die erste Fallanalyse den Blick auf die (misslingende) Synchronisierung verbaler und nonverbaler Modalitäten der Kommunikation gerichtet und herausgestellt, dass und wie sich im vorliegenden Fall Nadines Körper und Verbalität unabhängig voneinander machen. Nicht nur zeigt sich somit, wie mithilfe einer multimodalen Brille systematisch Probleme der Mensch-Roboter-Interaktion rekonstruiert werden können, sondern darüber hinaus werden auch Probleme der Integration der verschiedenen Systeme von Nadine deutlich, die sich in Interaktion eben nicht als personale *Einheit*, sondern als *Vielheit* präsentiert und damit verbunden nicht als vollwertige Kommunikationspartnerin reüssieren kann.

Demgegenüber richtete sich die Analyseperspektive der zweiten Fallanalyse auf die Beobachtung des Zusammenspiels von innenweltlichen und kommunikativen (hier: sprachbasierten) Prozessen. In diesem Kontext wurde dargelegt, dass und wie sich der Agent MAX als Kontrollfreak outet, der Usereingaben entlang starrer Pläne und Regeln beobachtet und abarbeitet. Dabei benötigt der ECA Schlagworte seines Gegenübers, um fortfahren zu können, also in das Fahrwasser bekannter Dialogprogramme (z.B. – wie hier – den Namen herausfinden) zu kommen. Im Zuge dieser Vereindeutlichung oder Verhärtung von mitunter unwägbarer sozialer Komplexität in klare Kausalschemata, konkurrieren unterschiedliche Interpretationen um einzelne Funktionen für Anschlussbeiträge, die letztlich in der Kommunikation Störungen verursachen. Gleichzeitig – und hier deuten sich Möglichkeiten an, Begegnungen von Menschen und Robotern/Agenten als wechselseitige Kontrollprojekte zu beobachten – versucht der/die Nutzer\*in den Agenten unterschwellig mit Schlagworten zu füttern, die von seinen Handlungsprogrammen verarbeitet werden können. Inwieweit ein solches Vorwegnehmen der innenweltlichen Prozesse des künstlichen Kommunikationsteilnehmers<sup>14</sup> tatsächlich Programm hat, bleibt in weiteren Untersuchungen abzuwarten.

Mit Blick auf die Gemeinsamkeiten unserer Gegenstände entsteht der Eindruck, dass beide Systeme (Nadine und MAX) auf je unterschiedliche Art und Weise

---

<sup>14</sup> Und damit der Versuch, den Kontrollfreak unter Kontrolle zu bringen.

Probleme produzieren, die auf fehlende oder für Kommunikationszwecke nicht ausreichende Synchronisation ihrer Komponenten zurückzuführen sind. Rund um Nadine sind die Module der Wahrnehmung so gestreut, dass nur in Momenten der Synchronisation aller Sensoren so etwas wie Signalisierung uneingeschränkter Kommunikationsfähigkeit entsteht. Dagegen liegen die Probleme bei MAX in der Verkopplung der Module in seiner Innenwelt. So scheitert die Kognitionsarchitektur von MAX gerade daran, dass ebenso verteilte Interpretationsmodule („first input“ und „discourse greeting friend“) nicht synchronisiert, sondern ausschusslogisch im Sinne einer entweder/oder-Entscheidung mit Antwortphrasen verknüpft werden. Folgestudien könnten nun – entsprechende Verfügbarkeit von Daten vorausgesetzt – versuchen, beide Analyseschwerpunkte zu kombinieren, um auf diese Weise genauer das Zusammenspiel von Oberfläche (das sichtbare Interface) und Tiefe (die Programmierung) und deren Relevanz für kommunikative Prozesse zu erhellen.

## Literaturverzeichnis

- Alač, Morana (2016): Social robots: Things or agents? In: *AI & Society* 31 (4), S. 519-535.
- Bergmann, Jörg R. (1982): Schweigephasen im Gespräch - Aspekte ihrer interaktiven Organisation. In: Hans-Georg Soeffner (Hg.): *Beiträge zu einer empirischen Sprachsoziologie*. Tübingen: Narr, S. 143-184.
- Bergmann, Jörg R. (1987): *Klatsch. Zur Sozialform der diskreten Indiskretion*. Berlin/New York: de Gruyter.
- Breazeal, Cynthia L. (2002): *Designing sociable robots*. Cambridge / MA: MIT Press.
- Cassel, Justine; Sullivan, Joseph W.; Prevost, Scott; Churchill, Elisabeth F. (Hg.) (2000): *Embodied Conversational Agents*. Cambridge; Massachusetts; London: MIT Press.
- Chomsky, Noam (1975): *Syntactic Structures*. Berlin: de Gruyter.
- Coulter, Jeff; Sharrock, W. (2007): *Brain, mind and human behavior in contemporary cognitive science: Critical assessments of the philosophy of psychology*. Ceredigion, UK: The Edwin Mellen Press.
- Edelman, Shimon (2008): *Computing the Mind: How the Mind Really Works*. Oxford, UK: Oxford University Press.
- Fuchs, Peter (1997): Adressabilität als Grundbegriff der soziologischen Systemtheorie. In: *Soziale Systeme* 3 (1997), S. 57-79.
- Garfinkel, Harold (1963): A conception of, and experiments with, „trust“ as a condition of stable concerted actions. In: O.J. Harvey (Hg.): *Motivation and social interaction: cognitive approaches*. New York: Ronald Press, S. 187-238.
- Garfinkel, Harold (1967): *Studies in Ethnomethodology*. Englewood Cliffs, N.J.: Prentice Hall.
- Garfinkel, Harold; Wieder, D. Lawrence (1992): Two incommensurable, asymmetrically alternate technologies of social analysis. In: Graham Watson; Robert M. Seiler (Hg.): *Text in context: studies in ethnomethodology*. Newbury Park, etc.: Sage, S. 175-206.
- Gesellensetter, Lars (2004): *Ein planbasiertes Dialogsystem für einen multimodalen Agenten mit Präsentationsfähigkeit*. Diplomarbeit, eingereicht an der Technischen Fakultät der Universität Bielefeld.
- Goffman, Erving (2008): *Rahmen-Analyse. Ein Versuch über die Organisation von Alltagserfahrungen*. Frankfurt am Main: Suhrkamp.
- Huber, Marcus (1999): JAM: A BDI-Theoretic Mobile Agent Architecture. In: *Proceedings of the Conference on Autonomous Agents*, S. 236-243.

Jefferson, Gail (1989): Preliminary notes on a possible metric which provides for a 'standard maximum' silence of approximately one second in conversation. In: Derek Roger; Peter Bull (Hg.): *Conversation: An interdisciplinary Perspective*. Clevedon, UK: Multilingual Matters, S. 166-195.

Kieserling, André (1999): *Kommunikation unter Anwesenden. Studien über Interaktionssysteme*. Frankfurt / M.: Suhrkamp Verlag.

Knoblauch, Hubert (2013): *Videographie. Einführung in die interpretative Videoanalyse sozialer Situationen*. Wiesbaden: VS Verlag.

Kopp, Stefan; Gesellensetter, Lars; Krämer, Nicole; Wachsmuth, Ipke (2005): A Conversational Agent as Museum Guide – Design and Evaluation of a Real-World Application. In: Themis Panayiotopoulos; Jonathan Gratch; Ruth Aylett; Daniel Ballin; Patrick Olivier; Thomas Rist (Hg.): *Intelligent Virtual Agents*. Berlin; Heidelberg: Springer, S. 329-343.

Krummheuer, Anita (2010): *Interaktion mit virtuellen Agenten? Zur Aneignung eines ungewohnten Artefakts*. Stuttgart: Lucius & Lucius.

Levinson, Stephen C. (2000): *Pragmatik. Konzepte der Sprach- und Literaturwissenschaft*. Band 39. Tübingen: Niemeyer Verlag.

Luhmann, Niklas (2002): *Einführung in die Systemtheorie*. Heidelberg: Carl-Auer Verlag.

Luhmann, Niklas (2008): Was ist Kommunikation? In: Ders.: *Soziologische Aufklärung 6. Die Soziologie und der Mensch*. Wiesbaden: VS Verlag, S. 109-121.

Magenat Thalmann, Nadia; Yuan, Junsong; Thalmann, Daniel; You, Bum-Jae (Hg.) (2016): *Context Aware Human-Robot and Human-Agent Interaction*. Singapore: Springer Singapore.

Magenat Thalmann, Nadia; Tian, Li; Yao, Fupin (2017): *Nadine: A Social Robot that Can*

*Localize Objects and Grasp Them in a Human Way*. In: S.R.S Prabaharan; Nadia Magenat

Thalmann; V. S. Kanchana Bhaaskaran, (Hg.): *Frontiers in Electronic Technologies*. Singapore: Springer Singapore, S. 1-23.

Meister, Martin (2011): Mensch-Technik-Interaktivität mit Servicerobotern. Ansatzpunkte für eine techniksoziologisch informierte TA der Robotik. In: *Technikfolgenabschätzung-Theorie und Praxis* 20 (1): S. 46-52.

Mondada, Lorenza; Schmitt, Reinhold (2010): Zur Multimodalität von Situationseröffnungen. In: Dies. (Hg.): *Situationseröffnungen. Zur multimodalen Herstellung fokussierter Interaktion*. Tübingen: Narr, S. 7-52.

Muhle, Florian (2013): Grenzen der Akteursfähigkeit. Die Beteiligung „verkörperter Agenten“ an virtuellen Kommunikationsprozessen. Wiesbaden: VS Verlag.

Muhle, Florian (2016): „Are you human?“. Plädoyer für eine kommunikationstheoretische Fundierung interpretativer Forschung an den Grenzen des Sozialen. In: Forum Qualitative Sozialforschung / Forum: Qualitative Social Research 17 (1).

Muhle Florian (2018a): Sozialität von und mit Robotern? Drei soziologische Antworten und eine kommunikationstheoretische Alternative. In: Zeitschrift für Soziologie 47 (3), S. 147-163.

Muhle, Florian (2018b): Begegnungen mit Nadine. Probleme der 'Interaktion' mit einem humanoiden Roboter. In: Angelika Pofert & Michaela Pfadenhauer (Hg.): Wissensrelationen. Beiträge und Debatten zum 2. Sektionskongress der Wissenssoziologie. Weinheim: Beltz Juventa, S. 499-511.

Pfeiffer, Thies; Liguda, C.; Wachsmuth, Ipke; Stein, Stefan (2011): Living with a Virtual Agent: Seven Years with an Embodied Conversational Agent at the Heinz Nixdorf MuseumsForum. In: Sara Barbieri, Katherine Scott und Luigina Ciolfi (Hg.): Proceedings of the International Conference Re-Thinking Technology in Museums 2011 - Emerging Experiences. Limerick: thinkk creative & the University of Limerick, S. 121-131.

Pitsch, Karola (2015): Ko-Konstruktionen in der Mensch-Roboter-Interaktion: Kontingenz, Erwartungen und Routinen in der Eröffnung. In: Ulrich Dausendschön-Gay; Elisabeth Gülich; Ulrich Krafft (Hg.): Ko-Konstruktion in der Interaktion: Die gemeinsame Arbeit an Äußerungen und anderen sozialen Ereignissen. Bielefeld: transcript, S. 229 – 258.

Schegloff, Emanuel A. (1991): Reflections on talk and social structure. In: Talk and Structure. Cambridge, UK: Polity Press, S. 44-70.

Schmitt, Reinhold (2015): Positionspapier: Multimodale Interaktionsanalyse. In: Ulrich Dausendschön-Gay; Elisabeth Gülich; Ulrich Krafft (Hg.). Ko-Konstruktionen in der Interaktion: Die gemeinsame Arbeit an Äußerungen und anderen sozialen Ereignissen. Bielefeld: transcript, S. 43-51.

Schneider, Wolfgang Ludwig (2008): Systemtheorie und sequenzanalytische Forschungsmethoden. In: Herbert Kalthoff; Stefan Hirschauer; Gesa Lindemann (Hg.): Theoretische Empirie. Zur Relevanz qualitativer Forschung. Frankfurt am Main: Suhrkamp. S. 129- 162.

Sidnell, Jack (2010): Conversation analysis. An introduction. Chichester, etc.: Wiley-Blackwell.

Suchman, Lucy (1987): Plans and situated actions. The problem of human-machine communication. Cambridge: Cambridge University Press.

Wachsmuth, Ipke (2010): "Ich, Max" – Kommunikation mit Künstlicher Intelligenz. In: Sutter, Tilmann; Mehler, Alexander (Hg.): Medienwandel als Wandel von Interaktionsformen. VS Verlag für Sozialwissenschaften, S. 135-157.

Zhao, Shanyang (2006): Humanoid social robots as a medium of communication. In: New Media & Society 8 (3), S. 401-419.

**Abbildungsverzeichnis:**

Abb. 1: Grafik aus eigenem Datenbestand

Abb. 2: Grafik aus eigenem Datenbestand

Abb. 3: Grafik aus eigenem Datenbestand

Abb. 4: Grafik aus eigenem Datenbestand

Abb. 5: Grafik aus eigenem Datenbestand

Abb. 6: Grafik aus eigenem Datenbestand

Abb. 7: Kopp et al. 2005: 333