

Ipke WACHSMUTH, Bielefeld

Kommunikation und Körper (Embodied Communication)

Einführung – Worum geht es?

Kommunikation – zweifelsohne eine der komplexesten und wichtigsten Fähigkeiten des Menschen – besteht längst nicht nur in der Weitergabe verbaler Informationen¹. Mehr als 65 Prozent des Austausches in einem Gespräch laufen vielmehr durch andere Kanäle: Gesten, Körperhaltung, Mimik, Sprachmelodie (Argyle, 1988). Bis zu 90 Prozent des Sprechens im natürlichen Diskurs wird durch Gesten begleitet (Nobe, 2000). Gestützt auf eine Vielzahl neuer detaillierter Befunde findet der Beitrag körperlicher Signale – speziell in der Kommunikation von Angesicht zu Angesicht – aktuelles und wachsendes theoretisches Interesse in der Sprachforschung und der Kommunikationsforschung im allgemeinen.

Nach Ansicht vieler Forscher entstand Sprache in Verbindung mit praktischem, also körperlichem, Handeln, ist demnach ohne Körper gar nicht vorstellbar. Und so verwundert es kaum, dass Sprechen immer von körperlichen Äußerungen begleitet ist, erst durch den Körper vollständig wird. Während wir sprechen, können wir z.B. die Form eines Objekts mit ein paar Handschlägen andeuten, durch Zeigen oder Blicke die Aufmerksamkeit auf ein Bezugsobjekt lenken oder Gesagtes durch emotionale Gesichtsausdrücke modifizieren. Auf diese Weise werden multimodal enkodierte, stark kontextabhängige Bedeutungen vermittelt.

Dies gilt auch für Aspekte des Verstehens. Zum Beispiel sind Gesten sensitiv für die Gesten anderer, sie koppeln Sprecher und Zuhörer in einer engen „sozialen Schleife“, in der sich nicht nur ihre Gesten, sondern auch ihre Gedanken wechselseitig beeinflussen. Kommunikation, so zeigt sich auch hier, besteht nicht nur darin, Wörter zu formulieren, sie ist eine Aktivität des ganzen Körpers. Der Körper bildet so allem Anschein nach auch die erste und grundlegendste Brücke in der sozialen Interaktion.

¹ Anfang der 1980er Jahre arbeitete der Autor in den USA mit Schülern, um zu beobachten, wie sie Mathematik lernen. Sie demonstrierten die Lösung von Aufgaben mit Klötzchen und anderem didaktischem Material, dabei zeigten sie hierhin und dorthin und beschrieben Formen und Lagen mit den Händen. Das hier ursprünglich geweckte Interesse an sprachbegleitenden Gesten wurde später zu einem Hauptforschungsthema, jetzt im Kontext natürlicherer Formen der Mensch-Maschine-Kommunikation.

Abgeleitet von dem lateinischen *communicari* („teilen, mitteilen“) wird Kommunikation als Informationsaustausch zwischen Personen verstanden, oder meist etwas spezieller als wechselseitige, weitgehend beabsichtigte Informationsübertragung zwischen mindestens zwei Partnern. Die verbreitete, auf die Informationstheorie von Shannon und Weaver zurückgehende Vorstellung, dass dabei ein Sender Informationen verschlüsselt, die ein Empfänger aufnimmt und entschlüsselt, kann die wechselseitige enge Partnerkopplung in menschlichen Kommunikationsprozessen nicht angemessen beschreiben. Mit der Betonung von symbolischer Informationsübermittlung vernachlässigen klassische Kommunikationsmodelle den Beitrag nicht-symbolischer Qualitäten, die speziell in der Kommunikation von Angesicht zu Angesicht wirksam sind.



Abb. 1 (aus Sowa & Wachsmuth, 2002)

Über Wortsymbole der Sprache hinaus werden Bedeutungen vermittelt, die nicht auf einem konventionalisierten Zeichencode basieren, aber dennoch verständlich sind. Durch ikonische Gestik kann beispielsweise ein Vorstellungsbild in verkörperter Form extern repräsentiert und übermittelt werden (McNeill, 1992), wie in Abb. 1 mit der Beschreibung einer Containerform gezeigt. Ein solches gestisches Zeichen trägt die Bedeutung durch bildliche Ähnlichkeit zum vorgestellten Bezugsobjekt in sich. Des Weiteren kann der Ausdruck in Gesicht und Stimme emotionale Zustände übermitteln, die in ihrer Subtilität kaum durch Verbalisierungen mitteilbar sind.

Mit dem Begriff „verkörperte Kommunikation“ (*embodied communication*) soll der entscheidende Beitrag körperlicher Signale in der Kommunikation thematisiert werden. Diese Sicht impliziert, dass ein vollständigeres und korrekteres Verständnis von Kommunikation – über die formale Untersuchung natürlicher Sprache hinaus – systematisch den Einbezug physischer und biologischer Ressourcen verlangt. Eine Herausforderung, die sich an eine solche neue Konzeption von Kommunikation stellt, ist der Entwurf operationaler Modelle, die theoretisch begründet und empirisch geleitet sind und die spezifizieren, wie mentale Prozesse und Verkörperung in der Kommunikation zusammenwirken.

Maschinenkommunikation – Künstliche humanoide Agenten

Eine Vielzahl jüngerer Arbeiten in der Künstlichen Intelligenz und Robotik greift die Fragestellung verkörperter Kommunikation in technischem Zusammenhang auf. Speziell widmen sich die Arbeitsrichtungen von „embodied conversational agents“ (Cassell et al., 2000) und „virtual humans“ (Gratch et al., 2002) dem Ziel, Schlüsselaspekte der Verkörperung in der Kommunikation zu modellieren und auf der anderen Seite anthropomorphe Assistenzsysteme mit zahlreichen Anwendungsperspektiven zu erarbeiten.

Mit dem künstlichen Agenten „Max“, den wir in mehrjähriger Arbeit im Sonderforschungsbereich 360 „Situierete Künstliche Kommunikatoren“ (Rickheit & Wachsmuth, 1996) an der Universität Bielefeld entwickelt haben, untersuchen wir „verkörperte Kommunikation“ mit Computersimulationen in virtueller Realität (Kopp et al., 2003). Mit synthetischer Stimme und einem animierten Körper kann Max sprechen, gestikulieren und Gesichtsausdrücke zeigen. Über Mikrofon und ein Tracking-System kann Max sein Gegenüber auch „hören“ und „sehen“ und Sprache, Gestik und Blickrichtung des Menschen als Eingaben verarbeiten.



Abb. 2 Dialogsituation mit Max

In unserem Forschungsszenario geht es um das Bauen von Objekten, zum Beispiel eines Flugzeuges, aus einem *Baufix*-Konstruktionsbaukasten. Hieran wird erprobt, ob Max sich in wechselnden Situationen soweit „verständlich“ erweist, dass er im Dialog mit einem Menschen standhält.²

² Weil unser Agent sich einerseits multimodal (mit Sprache, Gestik und auch Gesichtsmimik) äußern kann und er sich andererseits mit der Assemblierung virtueller Objekte auskennt, wurde er auf MAX – für „Multimodaler AssemblierungseXperte“ – getauft.

Mensch und Max stehen sich dabei an einem Tisch gegenüber (siehe Abb. 2; mit Ausnahme des Menschen ist die dargestellte Szene eine projizierte virtuelle Realität). Auf dem Tisch liegen verschiedene Bausteine, Leisten mit drei oder fünf Löchern, Schrauben mit farbigen Köpfen, ein Würfel mit Löchern auf allen Seiten etc., die im Verlauf des Dialogs zusammengebaut werden. Sowohl Mensch als auch Max können durch natürlichsprachliche Instruktionen und Gesten den Zusammenbau einzelner Teile veranlassen, der in Simulation, unterlegt durch realistische Geräusche, ausgeführt wird. Das läuft zum Beispiel wie folgt ab:

- Max: Hallo, ich bin Max, was kann ich für Sie tun? [*lächelt freundlich*]
Mensch: Erkläre, wie man einen Propeller baut.
Max: Lass es uns zusammen machen. Zuerst muss eine Schraube in die Mitte einer Dreilochleiste gesteckt werden.
Mensch: Steck die gelbe Schraube in die Leiste. [*zeigt auf eine der Dreilochleisten; die Schraube springt ins erste Loch*]
Max: Nein, das war das falsche Loch. [*die Schraube rutscht wieder heraus*]
Max: Ich erkläre es dir noch einmal. Eine Schraube muss in die Mitte einer Dreilochleiste gesteckt werden.
Mensch: Steck die Schraube in die Mitte der Dreilochleiste. [*die Schraube springt nun in die Mitte der zuvor gewählten Dreilochleiste*]
Max: Okay, richtig gemacht. Jetzt stecke das hier in eine Dreilochleiste. [*zeigt auf die Leiste mit der Schraube*]
Mensch: Mach du das.
Max: Okay, dann übernehme ich das. [*die Objekte werden entsprechend verbunden*]
Max: Und jetzt müssen die Leisten quer zueinander gedreht werden. [*dreht seine Hände gegeneinander*] usw.

Die sprachlichen Äußerungen von Max werden, unter Anpassung von Parametern an die aktuelle Situation und inklusive der Generierung passender Gesten, aus einem Repertoire stereotyper Aussageformen erzeugt (Kopp & Wachsmuth, 2004). Mit simulierten Gesichtsmuskeln kann Max dabei auch „emotionale Zustände“ zum Ausdruck bringen, die unter anderem von dem Erreichen oder Misslingen kommunikativer Ziele beeinflusst werden.

Neben dem Aspekt der technischen Machbarkeit sind unsere Forschungsarbeiten auch mit der Erwartung verbunden, durch die Entwicklung und den Test konkreter Modelle neue Erkenntnisse über das Funktionieren menschlicher Kommunikation zu gewinnen. Wie funktioniert beispielsweise das zeitliche Zusammenspiel von Sprechen und Zeigen? Wie wird das

Abwechseln im Dialog gesteuert? Mit einer starken Verzahnung von linguistischen, psycholinguistischen und informatischen Forschungsmethoden bietet der DFG-Sonderforschungsbereich 360 hierfür ein hervorragendes Umfeld.

In aktuellen Arbeiten wollen unter anderem herausfinden, was es erfordert, formbeschreibende Gesten imitieren und verstehen zu können (Kopp et al., 2004), und welche Rolle Emotionen in der Kommunikation spielen (Becker et al., 2004). Gleichzeitig entwickeln wir – in enger Verbindung von erkenntnisgeleiteter Forschung und Anwendungserprobung – mit Max neuartige Formen der Mensch-Maschine-Kommunikation.



Abb. 3 Max im HNF

Als erste Anwendung fungiert Max seit Januar 2004 im Heinz Nixdorf MuseumsForum (HNF) in Paderborn als Auskunftgeber, der Besuchern in Dialogen Wissenswertes über einen der Ausstellungsbereiche und über das Computermuseum erklärt (Abb. 3). Dabei kann Max zwischen mehreren Menschen hin und her schauen (die er eigentlich nur daran vermutet, dass sein Kameraauge hautfarbene Flecke von Gesichtsgröße wahrnimmt); er kann sowohl in den virtuellen Raum (z.B. auf sich) zeigen als auch auf Exponate im Realraum, deren Koordinaten ihm bekannt sind, und simuliertes emotionales Verhalten an den Tag legen.

In der Museumsumgebung sammeln wir erste (positive) Erkenntnisse über die Akzeptanz, die „Normalmenschen“ einem künstlichen Ansprechpartner entgegen bringen; eine systematische Studie wird derzeit vorbereitet.

Ausblick – Wie geht es weiter?

Das Paradigma der Verkörperung ist derzeit eins der viel versprechendsten in der Kognitionswissenschaft. Mit dem vorliegenden Beitrag sollte die Forschungsrelevanz dieses Paradigmas auch – und gerade – für die Kommunikation verdeutlicht werden. Eine der Erwartungen ist es, dass die Konstruktion und experimentelle Untersuchung „künstlicher Kommunikatoren“ dazu beiträgt, ein tieferes Verständnis verkörperter Kommunikation zu erlangen.

Parallel dazu untersuchen Forscher in aller Welt die Hirnmechanismen zur Handlungskoordination zwischen Sozialpartnern. Ihre jüngsten Entdeckungen haben Einfluss auf die Forschung in der Künstlichen Intelligenz, der humanoiden Robotik und der Mensch-Maschine-Kommunikation, und sie werden umgekehrt davon beeinflusst.

Ausgehend u.a. von der Entdeckung der Spiegelneuronen in der Forschung an nichtmenschlichen Primaten (Gallese et al., 1996) wird heute angenommen, dass auch bei Menschen eine enge sensomotorische Kopplung in Wahrnehmung und Handlung besteht, die uns körperliche Äußerungen des Gegenübers durch eine Art Resonanz im eigenen Körpermodell verstehen lässt (Sebanz et al., 2003). Bei rapide zunehmenden Einsichten fehlt es jedoch an einer integrativen Sicht, die sensomotorische Modelle und ihre Rolle im „social loop“ auf höherer Ebene mit funktionalen Modellen kommunikativer Mechanismen zusammen führen könnte. Eine vollständigere Untersuchung dieser Phänomene berührt viele weitere Aspekte, und viele Disziplinen.

In umfassender Weise ist deshalb „verkörperte Kommunikation“ (bei Mensch und Maschine) das Thema eines Forschungsjahres am Zentrum für interdisziplinäre Forschung (ZiF) der Universität Bielefeld, das ab Oktober 2005 international führende Vertreter der kognitiven, der Neuro- und der Computerwissenschaften zusammen bringen soll (Wachsmuth & Knoblich, 2005). Der Fokus der geplanten Arbeiten richtet sich dabei sowohl auf das bessere Verständnis der menschlichen Kommunikation und ihrer Evolutionsgeschichte als auch auf das Ziel, zukünftig einmal mit Robotern und anderen Maschinen auf natürliche Weise kommunizieren zu können.

Das Gesamtziel des ZiF-Projektes ist es, aus zahlreichen Einzelerkenntnissen der beteiligten Disziplinen eine integrierte Sicht körperlich verankerter Kommunikation zu entwickeln, von der eine Vielzahl an theoretischen, empirischen und anwendungsbezogenen Fragestellungen wechselseitig profitieren kann.

Literatur

- Argyle, M. (1988). *Bodily Communication (2nd edition)*. New York: Methuen & Co.
- Becker, C., Kopp, S., & Wachsmuth, I. (2004). Simulating the emotion dynamics of a multimodal conversational agent. In E. André et al. (Eds.): *Affective Dialogue Systems (pp. 154-165)*. Berlin: Springer (LNAI 3068).
- Cassell, J., Sullivan, J., Prevost, S., & Churchill, E. (eds.) (2000). *Embodied Conversational Agents*. Cambridge, MA: MIT Press.
- Gallese, V., Fadiga, L., Fogassi, L., & Rizzolatti, G. (1996). Action recognition in the premotor cortex. *Brain*, 119, 593–609.
- Gratch, J., Rickel, J., André, E., Badler, N., Cassell, J., & Petajan, E. (2002). Creating Interactive Virtual Humans: Some Assembly Required. *IEEE Intelligent Systems*, July/August 2002 (2 1094-7167/02).
- Kopp, S., Jung, B., Lessmann, N., & Wachsmuth, I. (2003). Max – a multimodal assistant in virtual reality construction. In: *KI - Künstliche Intelligenz 4/03*, 11-17.
- Kopp, S., Sowa, T., & Wachsmuth, I. (2004). Imitation games with an artificial agent: from mimicking to understanding shape-related iconic gestures. In A. Camurri & G. Volpe (Eds.), *Gesture-based Communication in Human-Computer Interaction (pp. 436-447)*. Berlin: Springer (LNAI 2915).
- Kopp, S., & Wachsmuth, I. (2004). Synthesizing multimodal utterances for conversational agents. *Journal of Computer Animation and Virtual Worlds*, 15, 39-52.
- McNeill, D. (1992). *Hand and mind: What gestures reveal about thought*. Chicago: University of Chicago Press.
- Nobe, S. (2000). Where do most spontaneous representational gestures actually occur with respect to speech? In McNeill, D. (ed.), *Language and Gesture (pp. 186-198)*, Cambridge University Press.
- Rickheit, G., & Wachsmuth, I. (1996). Collaborative Research Centre "Situating Artificial Communicators" at the University of Bielefeld, Germany. *Artificial Intelligence Review* 10, 165-170.
- Sebanz, N., Knoblich, G., & Prinz, W. (2003). Representing others' actions: Just like one's own? *Cognition*, 88, B11-B21.
- Sowa, T., & Wachsmuth, I. (2002). Interpretation of Shape-Related Iconic Gestures in Virtual Environments. In I. Wachsmuth & T. Sowa (Eds.): *Gesture and Sign Language in Human-Computer Interaction (pp. 21-33)*. Berlin: Springer (LNAI 2298).
- Wachsmuth, I., & Knoblich, G. (2005). "Verkörperte Kommunikation bei Menschen und Maschinen" <http://www.uni-bielefeld.de/ZIF/FG/2005Communication/>