

Multimediale Modelle multimodaler Kommunikation Motion-Capturing in der computergestützten Gestenforschung

Schüller, Daniel

schueller@humtec.rwth-aachen.de
RWTH Aachen, Deutschland

Mittelberg, Irene

mittelberg@humtec.rwth-aachen.de
RWTH Aachen, Deutschland

Unter redebegleitende Gestik werden hier kinetische Arm-, Hand-, und Kopfbewegungen und –konfigurationen verstanden, wie sie von Sprechenden und ihren zuhörenden und zuschauenden Dialogpartnern mehr oder weniger bewusst im Rahmen der mündlichen Face-to-Face-Kommunikation eingesetzt werden (vgl. Kendon 2004, Müller 1998, Müller et.al. 2013, 2014). Bei redebegleitenden Gesten handelt es sich stets um performative, temporale und spontane Vollzüge. Im Unterschied zur Lautsprache lassen sich zunächst keine expliziten syntaktischen und semantischen Regeln erkennen, die das Phänomen unterfüttern. Somit erscheint redebegleitende Gestik (im Unterschied zu sog. Emblemen) als idiosynkratischer, aber dennoch integraler und auch musterhaft unterfütternder Bestandteil von kommunikativer Interaktion vermittelt sog. *natürlicher* (körpereigener) *Medien*. Gesprochene Sprache wird somit als multimodaler Vollzug beschrieben, von dem sich die diskursintegrierten Gesten nicht im Sinne einer eigenen medialen Spur abtrennen lassen.

Neben den etablierten, überwiegend rein qualitativ arbeitenden Methoden zur Untersuchung redebegleitender Gestik entwickeln sich zunehmend computerbasierte Verfahren, welche zusätzlich zur qualitativen Analyse von Videomaterial auch quantitative, auf numerische Daten gestützte Analyseperspektiven eröffnen. Im Natural Media Lab der RWTH Aachen werden beispielsweise Gesprächspartner mittels eines markerbasierten, optischen 3D- *Motion-Capture* -Systems aufgenommen und aus dem Verbund von MoCap-, Video- und Audiodaten ein Korpus erstellt. Die multimediale Transkription (Jäger 2004, 2012) versetzt das Korpus in den theoretischen Status eines digitalen Modells (Schüller und Mittelberg, 2017), das für GestenforscherInnen im Sinne eines empirischen Relativs an die Stelle der (Summe der) realen,

raumzeitlichen Bewegungen der Sprechenden tritt und die verschiedenen Modalitäten (gesprochene Sprache, Gestik) der Kommunikationssituation erfasst.

Der Vorteil dieses in den letzten Jahren mit Informatikern der RWTH Aachen entwickelten Verfahrens gegenüber der klassischen Videoanalyse ergibt sich hierbei aus der zusätzlichen multimedialen Verschränkung der Annotationen des Gestenforschers mit den numerischen Daten des MoCap-Systems. Diese Methodik unterstützt und ergänzt die Analyse redebegleitender Gesten insofern, als diskursintegrierte Gestik nun auch probandenübergreifend und quantitativ untersucht werden kann, indem auf den MoCap-Daten ein digitaler Algorithmus als operative Instanz eines Ähnlichkeitsmodells arbeitet (Beecks et. al. 2015, 2016; Schüller et.al. 2017), der mittels Gesten-Signaturen algorithmisch nach Ähnlichkeiten im kinetischen Verhalten der Probanden sucht. Weiterhin eröffnet der notationale Charakter (Goodman 1997) des digitalen Modells die Möglichkeit zur Erstellung von Diagrammen (Schüller und Mittelberg 2016) zur Visualisierung quantitativer Daten wie Beschleunigung, Abstände von Händen und Armen zum Körper, Nutzung des Gestenraumes (McNeill 1992) etc., während der multimodalen Artikulation gesprochener Sprache. Drittens gibt es die Möglichkeit der Projektion von Bewegungsspuren auf die Videodaten, sodass auch eine visuelle Repräsentation der ansonsten nicht sichtbaren numerischen Daten erfolgt und die Videodaten anreichert.

Erkenntnis- und wissenschaftstheoretisch betrachtet ist eine Reflexion des theoretischen Status solcher Modelle hochinteressant, da es sich um mittels digitaler, technischer Verfahren erzeugte, zeichenbasierte Transkriptionen realer Ereignisse handelt, die letztlich den Untersuchungsgegenstand empirischer Forschung bilden und somit zentral an der Gegenstandskonstitution beteiligt sind (Jäger 2012). Welchen Einfluss hat die Anwendung verschiedener Abstraktionslevel auf die Gegenstandskonstitution? Welche Zeichenprozesse sind an diesem transkriptiven Verfahren beteiligt?

In unserer Posterpräsentation wollen wir zunächst den eigentlichen Forschungsgegenstand – multimodalen Sprachgebrauch in körpereigenen Medien am Beispiel des Verbundes aus gesprochener Sprache und redebegleitender Gestik – erfassen. Hierauf folgt eine zeichentheoretische Beschreibung der technischen, transkriptiven Verfahren der Korpuskompilation im Natural Media Lab der RWTH Aachen. Darauf aufbauend soll ein Fokus auf die wechselseitige multimediale Verschränkung von Annotationen, Sprach- und Videodaten, numerischen MoCap-Daten, Algorithmus, und den hier jeweils involvierten Abstraktionsleveln gelegt werden.

Bibliographie

Beecks, C. / Hassani, M. / Hinnell, J. / Schüller, D. / Brenger, B. / Mittelberg, I. / Seidl, T. (2015): *Spatiotemporal Similarity Search in 3D Motion Capture*

Gesture Streams, in: Proceedings of the 14th International Symposium on Spatial and Temporal Databases (SSTD), Seoul, South Korea, August 26-28, 2015. S.355–372.

Beecks, C. / Hassani, M. / Hinnell, J. / Schüller, D. / Brenger, B. / Mittelberg, I. / Seidl, T. (2016): *Efficient Query Processing in 3D Motion Capture Databases via Lower Bound Approximation of the Gesture Matching Distance*, in: International Journal of Semantic Computing, 10(1), 5-25.

Goodman, Nelson (1997): *Sprachen der Kunst. Entwurf einer Symboltheorie*. Frankfurt a. M.: Suhrkamp.

Jäger, Ludwig (2004): *Transkription - zu einem medialen Verfahren an den Schnittstellen des kulturellen Gedächtnisses*, in: TRANS Internet-Zeitschrift für Kulturwissenschaften, 15 Nr., September 2004.

Jäger, Ludwig (2012): *Transkription*, in: **Christina Bartz / Ludwig Jäger / Marcus Krause / Erika Linz (eds.):** *Handbuch der Mediologie. Signaturen des Medialen*. München: Fink, S.306-315.

McNeill, David (1992): *Hand and mind: What Gestures Reveal about Thought*. Chicago: Chicago University Press.

Müller, Cornelia (1998): *Redebegleitende Gesten. Kulturgeschichte - Theorie – Sprachvergleich*. Berlin: Berliner Wissenschafts-Verlag.

Müller, Cornelia / Alan Cienki / Ellen Fricke / Silva Ladwig / David McNeill / Sedinha Teßendorf (eds.) (2013): *Body – Language – Communication. An International Handbook on Multimodality in Human Interaction* (Handbooks of Linguistics and Communication Science 38.1). Berlin, Boston: Mouton de Gruyter.

Müller, Cornelia / Alan Cienki / Ellen Fricke / Silva Ladwig / David McNeill / Jana Bressemer (eds.) (2014): *Body – Language – Communication. An International Handbook on Multimodality in Human Interaction* (Handbooks of Linguistics and Communication Science 38.2). Berlin, Boston: Mouton de Gruyter.

Schüller, D. / Beecks, C. / Hassani, M. / Hinnell, J. / Brenger, B. / Seidl, T. / I. Mittelberg (2017): *Automated Pattern Analysis in Gesture Research: Similarity Measuring in 3D Motion Capture Models of Communicative Action*, in: Digital Humanities Quarterly, 2017:11.2.

Schüller, D. / Mittelberg, I. (2016/ erschienen 2018): *Diagramme von Gesten: Eine zeichentheoretische Analyse digitaler Bewegungsspuren*, in: Zeitschrift für Semiotik 38, 3-4, S.7-38.

Schüller, D. / Mittelberg, I. (2017/ erschienen 2018): *Motion-Capture-gestützte Gestenforschung. Zur Relevanz der Notationstheorie in den Digitalen Geisteswissenschaften*, in: Zeitschrift für Semiotik 39, 1-2, S.109-146.