

Individual differences in speakers' gesture spaces: Multi-angle views from a motion-capture study

Matthias A. Priesters (priesters@humtec.rwth-aachen.de)

Natural Media Lab, Human Technology Centre, RWTH Aachen University
Theaterplatz 14, 52062 Aachen, Germany

Irene Mittelberg (mittelberg@humtec.rwth-aachen.de)

Natural Media Lab, Human Technology Centre, RWTH Aachen University
Theaterplatz 14, 52062 Aachen, Germany

Abstract

The approach presented in this paper aims to contribute to an account of the three-dimensionality of gesture space. Here, gesture space is assumed to be dynamically constructed and adaptive to the communicative situation. Making use of an optical motion-capture system, volumetric representations of gesture spaces were generated, based on gesture data from semi-structured interviews with four participants. The data were coded according to gesture phases and the gestures' communicative functions. We compare speakers' gesture rates and spatial distribution of gestures, both of which vary strongly across speakers.

Keywords: gesture space; individual differences; motion capture; methodology

Introduction

Understanding how space is used for communicative purposes is a central concern in gesture research. A substantial body of research has provided valuable insights into how the space parameter factors into conceptual, structural, and interactive dimensions of form and meaning in both co-verbal gestures and signed languages. However, only a relatively small number of studies have focused on the specific distribution of certain types of gestures (i.e., beats, pointing, iconic, or metaphoric gestures) across particular regions of gesture space (e.g., Pedelty, 1987; McNeill, 1992). Individual differences regarding people's preferred use of gesture space also await more systematic studies (cf. Bergmann & Kopp, 2010). The motion-capture study presented in this paper aims to shed new light on these interrelated aspects by (a) comparing gesture rates across speakers and (b) deriving and visualizing spatial patterns consisting only of those points and segments of gesture space that a speaker actually traverses with her gesturing hands during a given discourse.

Designed as a 'proof of concept' for a novel data-driven approach to gesture space, the primary aim of the pilot study reported on here was to digitally visualize a speaker's use of gesture space (Priesters, 2012). The results reported in the present paper confirm that people's individual gesturing styles and gesture spaces differ in several respects, even though they may be speaking about similar topics. Besides great variation in terms of gesture rate, the global shapes and internal structures of people's gesture spaces show clearly pronounced idiosyncratic profiles.

Gesture spaces: Emergent, adaptive patterns of communicative bodily action

Studies examining the use of gesture space often rely on rectangular (in some cases cube-shaped) grids for transcribing the locations of gestures in relation to the speaker's body. The most widely-used gesture space transcription scheme was introduced by Pedelty (1987) and became a de facto standard through its adoption by McNeill (1992). Since the original scheme was developed for a specific genre and setting (i.e., narratives by single seated persons) it cannot – and was probably not meant to – capture all possible scenarios or spatial properties of gestural communication. In the present context, the inability of 'shallow disk'-style models to account for gesture spaces' depth dimension is crucial, especially when considering interaction-oriented gestures (Bavelas, 1994). Furthermore, the way gesture space typically is divided into specific zones may (mis)lead us to assume a preexisting static structure of gesture space, in which the distinction between *central* and *peripheral* zones has been instrumental. The present study was motivated by the question whether gesture space segments, such as center, should rather be determined based on observed patterns of use.

We understand individual and shared gesture spaces as dynamic, adaptive constructs (Mittelberg, 2010; Priesters, 2012). Assuming that "gesture space is the area defined by the positions of [the speakers'] hands in different points in space" (Rodrigues, 2010, 29), our view rests on three tenets: (1) We distinguish a speaker's *kinesphere* (Laban, 1966), i.e., the space surrounding a speaker that is available for gesturing, from her *gesture space*, i.e., those points and segments of space that are actually populated by gestures and recruited for meaningful communication. As the notion "sphere" suggests, kinesphere and gesture space are spherical in nature, centered on the speaker's body; due to human anatomy most movements of the extremities take on the form of arcs around certain joints (Laban, 1966; Martell, 2005). Describing gesture space in terms of squares or cubes would thus not adequately represent the dynamics and scope of human communicative actions. (2) We see speakers' gesture spaces as discourse-driven spatial structures emerging through their purposeful bodily (inter)actions during communication. As speakers usually prefer to gesture in certain parts of their kinespheres and thus do not use space homogeneously, gesture spaces

exhibit zones with varying degrees of density. (3) The spatial location of a gesture does not seem arbitrary, but motivated, e.g., it may be oriented towards certain elements in the on-going interaction and/or its environment (e.g., Goodwin, 2007). Sweetser and Sizemore (2008), for instance, argue that the location of gestures is influenced by their communicative purposes; in particular the authors distinguish representing gestures performed in ‘unclaimed’ space from turn-organizing gestures performed towards the interlocutor.

Study

In July 2012, we conducted a pilot study in the Natural Media Lab¹ at RWTH Aachen University with the goals of gathering sufficient data for the development of a motion-capture-based gesture space visualization method and of testing the qualitative finding by Sweetser and Sizemore (2008, 48), that “[g]esturers divide up space in functionally relevant [...] ways”.

Interviews

Nine semi-structured interviews with German native speakers about life and career decisions were conducted. During the interviews, which ranged from 10–20 minutes in length, the participants were seated within an optical motion-capture system with reflective markers applied to their hands, arms and body. To assess handedness, the questionnaire by Cohen (2008) was administered beforehand. Four participants (two male, two female, aged 25–32 years, all predominantly right-handed; Table 1) were selected for further analysis. Of the omitted five participants, two were excluded because no proper flow of conversation could be established, a third was too distracted by the adhesive markers applied to her hands. The remaining two participants were planned to be included in the study, but had to be sidelined for lack of time available for post-processing and coding.²

Coding

The video data from the interview sessions were annotated using ELAN³ by a single coder. The gesture data were segmented into gesture units and phases (Kendon, 2004; Kita, Gijn, & Hulst, 1998; McCleary & Leite, in press), which were coded independently for each hand, capturing the exact intervals during which each hand was engaged in gesturing. Only *expressive* gesture phases (i.e., *strokes* and *holds*; Kita et al., 1998) were considered further, since they are taken to contribute to a person’s gesture space built up during a given conversation. For each expressive phase, its dominant communicative function (Müller, 1998) was coded, according to a slightly expanded version of the functional category system (*referential*, *pragmatic*, *interactive* categories) proposed by Kendon (2004).

¹<http://www.humtec.rwth-aachen.de/nmlab>

²Coding of these data is in progress, so they will be included in future analyses.

³<http://tla.mpi.nl/tools/tla-tools/elan>

Processing and visualization

Based on the ELAN annotations, volumetric representations of the participants’ gesture spaces were computed in MATLAB for the entirety of the interviews and separately for each functional category. A participant’s kinesphere is represented by a three-dimensional matrix, each segment of which stands for one cubic centimeter of space. The annotated video files were synchronous with the motion-capture data, so the time codes of the annotations could be easily converted to motion-capture frame numbers. For all recorded motion-capture frames that were part of an expressive phase of the respective hand, the spatial segments occupied by the hand during the frame were marked. Summing up all these marks resulted in a matrix containing the density structure of the gesture space. Note that *spatial gesture density* is thus defined in terms of the total time that spatial segments were occupied by hands during expressive gesture phases, not in terms of gesture phases per segment. From the participants’ matrices, two types of visualizations were generated: One shows ‘heat maps’ of gesture spaces, condensed into two dimensions (Fig. 1).⁴ The other method, more suited to multimedia-based forms of presentation, displays 3D images of gesture spaces as ‘density clouds’ inside the participants’ kinespheres (not shown here; see Priesters, 2012).

Individual differences in gesture spaces

Gesture rates

The participants strongly differ regarding gesture rates. Both the numbers of expressive gesture phases and their cumulative durations were extracted from the ELAN annotations. As mentioned above, the lengths of the interviews varied, therefore the numbers of produced gestures and the durations were averaged per minute (Table 1). The gesture rates of P01 and P02 are in the same range, but only at a fraction of P03’s and P04’s rates. P04’s gesture rate is roughly twice as high as that of P03. Another difference is the degree to which the participants favored their dominant hand: Whereas P01 and especially P04 show a strong tendency towards their right hand, this tendency is weaker in P03 and ambiguous in P02, who performs more expressive phases with her right hand, but longer phases with her left hand. Interestingly, there appears to be no clear correlation between degree of handedness and preference of the dominant hand, especially when comparing P03 and P04.

Spatial distribution

P01’s gesture space (Fig. 1a) is densest near the hands’ rest positions. Overall, this gesture space has a significant tendency towards the right side. The area in front of the chest has a low density; the area above chest height is only sparsely filled with gesture strokes (there are only faint traces above

⁴Note that the coloring in the individual images is relative to the range of values of the respective participant, as indicated by the accompanying scales. The original set of images also includes a side view, which was omitted in this paper due to limits of space.

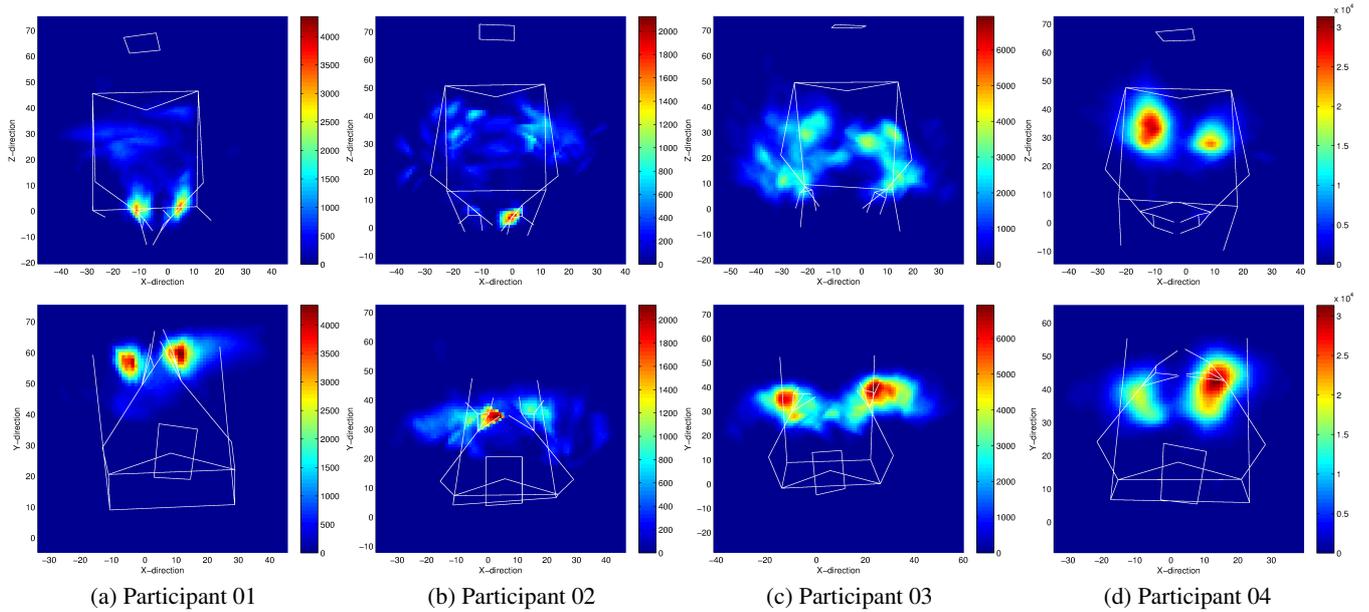


Figure 1: Gesture space density plots of four participants, frontal and top views, with body outlines.

Table 1: Numbers and duration in seconds of expressive gesture phases per minute, for both hands of each participant; degree of handedness (Cohen, 2008).

	count		duration		handedness decile
	LH	RH	LH	RH	
P01	2.97	4.35	1.04	1.75	3rd right
P02	3.12	3.45	1.56	1.37	3rd right
P03	11.46	13.65	6.55	8.74	7th right
P04	10.58	21.01	10.50	17.71	4th right

the right shoulder). As the view from above reveals, the right hand appears to have a larger movement radius than the left hand: the gesture space on the right side reaches further forward and outward.

P02’s gesture space (Fig. 1b) is very dispersed and fragmented, with a very low overall density, except for one hot spot around the left hand rest position, which is the result of long hold phases occurring in this location. The rest of the gesture space mostly consists of disjointed segments of low density distributed around the upper body and is somewhat more dense in front of the chest. From above, P02’s gesture space roughly has the shape of an arc around her sternum. She reaches comparatively far out to both sides, but apparently prefers the areas above the thighs for gesturing.

P03 shows the most expansive use of space of the four participants (Fig. 1c). His gesture space comprises various dense parts ranging from the rest positions to locations higher up in the area in front of the chest. These high-density areas are surrounded by extensive medium- and low-density regions occu-

pying a comparatively large part of the kinesphere. The right hand has a larger movement radius, reaching further out and up, almost to head height, whereas the left hand’s gestures show a higher spatial concentration, with the highest density in the lower chest region. Seen from above, P03’s gesture space has a clearly arced shape. There seem to be two distinct sub-spaces, one for each hand. Both are shaped like arcs around the elbows, which indicates that most gestures involved hand and forearm. These arcs are relatively thin in the depth dimension.

The structure of P04’s gesture space (Fig. 1d) is striking, particularly when compared to the other participants. Her gesture rate is by far the highest, but at the same time her gestures are spatially the most focused. P04 gestures almost exclusively in the upper chest region, with lower densities outside this area. There are regions of very low density around and above the rest positions, as well as above the right shoulder. The graph shows a large hot spot for the right hand, next to which we see a smaller and less dense spot for the left hand. This reflects the clear lateral preference that had already been indicated by the gesture rates. Viewed from above, P04’s gesture space is also condensed laterally, but broader in the depth dimension than for the other participants.

From this preliminary analysis, we can conclude that these findings attest to strong influences of people’s individual communicative styles on their gesture spaces, regarding both gesture rates and spatial distribution. Each speaker studied here has his/her unique spatial ‘signature’, so to speak.

Discussion and outlook

Our findings show a mixture of a certain systematicity and idiosyncratic patterns in gesture space structures, similar to

what Bergmann and Kopp (2010), who compared participants talking about the same referents, observed with regard to rate and handedness of iconic gestures. On the one hand, our results partially support McNeill's (1992, 304) observation that "[a]dult gestures are located in the space in front of the speaker, and the space is a flattened disk in which the horizontal and vertical dimensions are exploited, but the dimension of depth is truncated [...]." The majority of gestures are performed in front of the speakers' bodies. In the depth dimension, the gesture spaces are less extended than in the other dimensions (except for P04). Besides these rather broad systematicities, however, the finer-grained structures are highly idiosyncratic. While we have not discussed issues of functional patterns in the present paper, a more in-depth analysis has shown speakers' individual preferences for particular functions and also tendencies to distribute gestures with certain functions systematically across space (Priesters, 2012).

As our focus so far has been on developing an adequate methodology, we have not yet investigated the reasons for the noted variations. However, there are indications that some of the differences might be connected to personality traits such as extraversion (Hostetter & Potthoff, 2012) or the speakers' communicative skills (Hostetter & Alibali, 2007), though the small number of participants and the lack of data on these factors preclude reliable conclusions at this point.

The results of this pilot study support our understanding of gesture space as being adaptive, dynamically constructed and sphere-shaped. A re-evaluation of the traditional representations of gesture space, tracing back to McNeill's work, seems to be necessary. As for the data from the present study, more thorough qualitative analyses are needed to gain further insights into their idiosyncratic functional patterns. We further plan to enhance our measuring and visualization technique to also study the shapes and structures of shared gesture spaces. In view of interactively created spaces it will be particularly important to adequately represent their depth dimension. Another desideratum is the development of an annotation method that, while accounting for gesture spaces' three dimensions and their spherical shape, may also be applied to conventional video recordings.

Acknowledgments

This study was funded by the Excellence Initiative of the German Federal and State Governments. We thank the members of the Natural Media Group, especially Simon Harrison, Julius Hassemmer and Vito Evola, for valuable input at various stages of this project and Svetoslav Evtimov, Zhi Li and Brandon Montell for technical assistance during the study.

References

Bavelas, J. B. (1994). Gestures as part of speech: Methodological implications. *Research on Language and Social Interaction*, 27(3), 201–221.

Bergmann, K., & Kopp, S. (2010). Systematicity and idiosyncrasy in iconic gesture use: Empirical analysis and computational modeling. In S. Kopp & I. Wachsmuth (Eds.),

Gesture in embodied communication and human-computer interaction (pp. 182–194). Berlin/Heidelberg: Springer.

Cohen, M. S. (2008). *Handedness questionnaire*. Available from <http://www.brainmapping.org/shared/Edinburgh.php>

Goodwin, C. (2007). Environmentally coupled gestures. In S. D. Duncan, J. Cassell, & E. T. Levy (Eds.), *Gesture and the dynamic dimension of language* (pp. 195–212). Amsterdam/Philadelphia: Benjamins.

Hostetter, A. B., & Alibali, M. W. (2007). Raise your hand if you're spatial: Relations between verbal and spatial skills and gesture production. *Gesture*, 7(1), 73–95.

Hostetter, A. B., & Potthoff, A. L. (2012). Effects of personality and social situation on representational gesture production. *Gesture*, 12(1), 62–83.

Kendon, A. (2004). *Gesture: Visible action as utterance*. Cambridge: Cambridge University Press.

Kita, S., Gijn, I. van, & Hulst, H. van der. (1998). Movement phases in signs and co-speech gestures, and their transcription by human coders. In I. Wachsmuth & M. Fröhlich (Eds.), *Gesture and sign language in human-computer interaction* (pp. 23–35). Berlin/Heidelberg: Springer.

Laban, R. (1966). *Choreutics*. London: Macdonald & Evans.

Martell, C. (2005). *FORM: An experiment in the annotation of the kinematics of gesture*. Unpublished doctoral dissertation, University of Pennsylvania.

McCleary, L., & Leite, T. d. A. (in press). Turn-taking in Brazilian Sign Language: Evidence from overlap. *Journal of Interactional Research in Communication Disorders*.

McNeill, D. (1992). *Hand and mind: What gestures reveal about thought*. Chicago/London: University of Chicago Press.

Mittelberg, I. (2010). Geometric and image-schematic patterns in gesture space. In V. Evans & P. Chilton (Eds.), *Language, cognition and space* (pp. 351–385). London/Oakville: Equinox.

Müller, C. (1998). *Redebegleitende Gesten: Kulturgeschichte – Theorie – Sprachvergleich*. Berlin: Spitz.

Pedelty, L. L. (1987). *Gesture in aphasia*. Unpublished doctoral dissertation, University of Chicago.

Priesters, M. A. (2012). *Functional patterns in gesture space: Correlating location and function of gestures using 3D motion capture technology*. Unpublished master's thesis, RWTH Aachen University.

Rodrigues, I. G. (2010). Gesture space and gesture choreography in European Portuguese and African Portuguese interactions: A pilot study of two cases. In S. Kopp & I. Wachsmuth (Eds.), *Gesture in embodied communication and human-computer interaction* (pp. 23–33). Berlin/Heidelberg: Springer.

Sweetser, E., & Sizemore, M. (2008). Personal and interpersonal gesture spaces: Functional contrasts in language and gesture. In A. Tyler, Y. Kim, & M. Takada (Eds.), *Language in the context of use* (pp. 25–51). Berlin/New York: Mouton de Gruyter.