Towards a theory of Gesture Form Analysis
Imaginary forms as part of gesture conceptualisation, with empirical support from motion-capture data

A thesis in partial fulfilment of the requirements of
Degree of Doctor of Philosophy
in the Faculty of Philosophy
at the RWTH Aachen University
by

Julius Hassemer

Under the supervision of
Prof. Irene Mittelberg, Ph.D.
Prof. Dr. Klaus Willmes-von Hinckeldey
Prof. Dr. Cornelia Müller (European University Viadrina, Frankfurt (Oder))

Faculty of Philosophy
RWTH Aachen University

This thesis was defended on the 29th of April 2015
and is available in electronic format on the website of the university library.
Acknowledgements

I would like to thank my advisor, Irene Mittelberg, who, despite her numerous engagements, has supported me on every occasion, bearing with me writing a dissertation that wants to have the cake and eat it being just one example.

With you, Nicole Stein, I feverishly put together the first pieces of an idea on gesture form and you were also there for me in the eventful final stage of this dissertation.

I really did not know what I wanted to say until Leland McCleary told me how to say it.

Thanks, Gina Joue, for teaching me in scientific precision.

When I hung out with you, Vito Evola, I smoked more than I wanted, but learned more than I thought.

Thanks, Simon Harrison, for sharing some of your inexhaustible energy with me; this dissertation might have exhausted me otherwise.

Howdy Len Talmy, seldom someone is so forcefully critical and joltingly motivating at the same time.

Klaus Willmes, I hope some of your being a gentle all-rounder has rubbed off on this dissertation.

Thank you, Cornelia Müller, for accompanying me along so many important steps of my gesture journey, and especially for setting the tracks in the first place.

Thanks, Bela Brenger, Han Sloetjes, Marlon Meuters, Matthias Priesters, Svetoslav Evtimov and Zhi Li (in different order and combination) for programming the Elan Plugin for Automatic Annotation (EPAA); the motion-capture analysis tool in Java (MoJa); and the Data Extractor (DEx).

Dear participants of Study 1, Study 2 and pilots studies: thanks for your kind (and unremunerated) support, which is the soil of this dissertation.

Thank you, Mats Andrén, Jana Bressem, Suki Grandhi, Sarah Greifenstein, Ludwig Jäger, Philipp Krecklow, Silva Ladewig, Christina Schmitt, Sherman Wilcox, Bob Williams, Bodo Winter, Ulrike Wrobel, and others who have always pointed me in the right direction.
für Diana
Abstract

This dissertation presents the theory and application of Gesture Form Analysis, which systematises necessary processes of spatial conceptualisation for understanding a specific gesture type (pointing, drawing, representing etc.). Eight gesture form operations are proposed, which are combined in a modular fashion. They are applied in the analysis of gestures of the dissertation corpus and of other literature. The gesture type Pointing At Location is composed of the operations Articulator Profiling (e.g., straight extended arm, hand and index), Shape Profiling (the predominant axis, a line, is profiled), Continuation (the line is continued, indicating a direction), and Intersection (the line intersects with the surface of an object in a point: the location pointed at). The gesture type Drawing—while otherwise containing the same operations as Pointing At Location—requires adding the operation of Trace Leaving (by moving the arm, and hence all the other forms created by the above operations, the point leaves a trace in form of a line). The geometrical operations of Gesture Form Analysis (and the constituent forms specified in parentheses) consider specifically the spatial modality of gesture, taking a first step towards a strictly form-based gesture semantics. The systematic analysis of operations across gesture types results in a gesture typology, in which the dimensions of all constituent forms are defined; thus being open to computational application. The fundamental theoretical distinction of gesture form being non-equivalent to the mere physical form of the articulators (articulator form) is supported by an object-description study in an optical motion-capture system. The methods include operationalising three-dimensional gesture data, for example by a plugin for the video-annotation software Elan.
### Contents

Acknowledgements ii  
Abstract vi  
Contents vii  
List of figures, navigation xi  
List of abbreviations and coding conventions xix  
Introduction 1  
  Some gesture studies terminology 1  
  Examples of gesture classifications 2  
  The three parts of this dissertation 4  

**Part I: Aspects of form in gesture research, and the theoretical bases of gesture form** 7  
  1. Locating gesture form 8  
    1.1. Locating the research interest within the phenomenon of gesture 8  
    1.2. Two levels of gesture analysis 12  
    1.3. Three levels of gesture analysis, introducing ‘gesture form’ 13  
    1.4. Reasons for a gesture-specific methodology 16  
  2. Gesture form in different gesture types 19  
    2.1. ‘Form parameters’ describe articulator form, not gesture form 19  
    2.2. Gesture types: the hands imitating different practices 22  
    2.3. Iconic gestures can be more or less similar to their referent 26  
    2.4. Two different kinds of metonymy in gestures 27  
    2.5. Mandel’s “sketching” as a systematic analysis of gesture types 30  
    2.6. Conventionality, similarity, metonymy, and metaphor in orchestration 31  
  3. Forms in gesture literature 37  
    3.1. Observing bodily action 37  
    3.2. Pointing contains different forms 39  
      3.2.1. Line or vector 40  
      3.2.2. Point (intersection at a certain location) 42  
      3.2.3. Relation of forms in pointing 42  
    3.3. Drawing and pointing on common ground 44  
    3.4. Holding and moulding: surface-contact gestures 49  
    3.5. Representing in three dimensions 52  
      3.5.1. Mapping all or just selected form aspects 52  
      3.5.2. Different articulators represent 53  
      3.5.3. Representing and cognitive iconicity 53  
    3.6. Different kinds of form in a single gesture 55  
  4. Forms in perception 58  
    4.1. Abstraction 58  
    4.2. Surfaces in perception 61  
      4.2.1. Perception as active exploration 64  
      4.2.2. Visual perception 65  
      4.2.3. Tactile perception 69  
  5. Forms in language about space 73  
    5.1. Talmy’s topological dimensions 74  
    5.2. Profiling 76  
  6. Forms in signed languages 78  
    6.1. Sign is part of gesture is part of action 78  
    6.2. Blends in depicting scenes 79
Part II: Gesture Form Analysis as basis for a new gesture typology

7. Two principles of gesture form

7.1. Dimensions: Gesture form primitives

7.1.1. Spatial dimensions in gesture

7.1.2. Different form dimensions afford different states and actions in space

7.1.3. Summary of how different forms afford division

7.1.4. Summary of specific affordances in three-dimensional space

7.1.5. Fine-grained analysis for distinguishing spatial dimensions in gestures

7.2. Gesture form operations

7.2.1. Order of operations

7.2.2. Specific order: profiling operations in the beginning

7.2.3. The operations of Articulator and Shape Profiling

7.2.4. The operation of Continuation

7.2.5. The operation of Intersection

7.2.6. The operation of Adjacency (general)

7.2.7. The operation of Adjacency: Division

7.2.8. The operation of Adjacency: Enclosure

7.2.9. The operation of Trace Leaving

7.2.10. Alternative gesture form interpretations of a single gesture

7.2.11. Different articulator forms as ‘functionally homogeneous types’

7.2.12. Conceptually reducing the dimensions of gesture space

7.2.13. Where gesture form ‘begins’

7.2.14. Where gesture form ‘ends’: the resulting form

7.3. A definition of gesture form

7.4. Gesture form operations as a spatialisation of semiotic relations

8. Specific Gesture Form Analyses building up to a gesture typology

8.1. Study 1

8.2. Some differences to functional gesture categorisations

8.3. Developing a typology by systematic gesture analyses

8.4. Pointing and Drawing

8.4.1. The basis for Pointing and Drawing: Representing in one dimension

8.4.2. Directing

8.4.3. Pointing At Location (contact)

8.4.4. Pointing At Location (no contact)

8.4.5. Pointing At Object (contact)

8.4.6. Pointing At Object (no contact)

8.4.7. Drawing

8.5. Surface Contact Gestures: Holding and Moulding

8.5.1. The basis for Holding and Moulding: Representing in two dimensions

8.5.2. Holding

8.5.3. Moulding

8.6. Representing in three dimensions

8.6.1. An articulator representing itself in three dimensions

8.6.2. An articulator representing something else in three dimensions

8.6.3. Emblems

8.7. Other gesture types: complex pointing and ‘minimal content gestures’

8.7.1. Broad Pointing

8.7.2. Surface Directing

8.7.3. Line Sketching and Surface Sketching

8.7.4. Representing in zero dimensions

8.7.5. Indefinable Activity

8.8. A typology of basic gestural semantics based on Gesture Form Analysis
Part III: Quantified testing of gesture form: Measure and Shape Gestures

9. Design, setup and technical preparation of an object-description study (Study 2)

9.1. Experiences from Study 1 shaping the design of Study 2

9.1.1. Technical problems (system load)
9.1.2. Re-evaluating Data elicitation
9.1.3. Participants not interpreting one specific gesture at a time
9.1.4. Few gestures, long trials
9.1.5. Conclusion for a new study design

9.2. Study 2 in a nutshell

9.3. Participants

9.4. Lab setup

9.5. Technical setup

9.6. Minimal Marker Set

9.7. Object and video stimuli

9.8. Tasks

9.8.1. First part: object description task
9.8.2. Second part: observer task

9.9. Data post-processing

9.9.1. Hand Shape Coordinates
9.9.2. Motion-capture Java script (MoJa)
9.9.3. Elan Plugin for Automatic Annotation (EPAA)
9.9.4. Data Extractor (DEx)

9.10. Summarising the characteristics of Study 2

10. Are Measure and Shape Gestures in Study 2 perceived differently?

10.1. Motivation

10.2. Similarity of Measure and Shape Gestures

10.3. Gesture Form Analysis of Measure and Shape Gestures

10.3.1. Measure Gesture
10.3.2. Shape Gesture

10.4. Attaining Measure and Shape Gesture interpretations

10.4.1. Interpretation by participants
10.4.2. Interpretation by raters

10.5. Articulator form differences between Measure and Shape Gestures

10.5.1. Predicted differences
10.5.2. Operationalising Curve
10.5.3. Operationalising Curl

10.6. Hypothesis

10.7. The overall workflow of Study 2 including manual data selection

10.7.1. Excluding rest positions
10.7.2. Only index-and-thumb profile
10.7.3. Excluding gestures conveying spherical curvature
10.7.4. Excluding ambiguous gesture form
10.7.5. The temporal relation of low-movement phases and holds

10.8. Results

10.8.1. Difference between Measure and Shape Gestures confirmed
10.8.2. Categorical difference between Measure and Shape Gestures?

11. Discussion

11.1. Characteristics of Gesture Form Analysis
11.2. First thoughts on the integration into Talmy’s recent work
11.3. Two coarse speech analogies and the problem of categorical perception

12. Conclusion with suggestions for gesture form’s many loose ends

12.1. Potential follow-up studies

12.2. Gesture Form Analysis coding scheme

12.3. Contribution to computer-aided studies and human-computer interaction
List of figures, navigation

Since this dissertation heavily relies on graphical illustration, some conventions are to help quick navigation between figures:

1. The first number of the figure index shows the chapter to which the figure belongs.

2. Since all kinds of illustrations are called figures (including tables, designed lists) and thus numbered uniformly (instead counting up tables and figures independently), any figure number is informative about whether the desired figure is to be found above or below the figure at hand.

For the Gesture Form Analysis of examples from external literature, the references are not only in the caption, but also integrated in the picture in order to permanently link to the source of the image.

Figure 0.1: A gesture drawing a line (depicted in red) in the air

Figure 1.1: Three senses relevant in gesture perception.

Figure 1.2: 1.5 seconds of Elan coding of multiple articulators annotated in one tier each (Hassemer 2009:68).

Figure 1.3: Three levels of analysis including gesture form as an intermediate level.

Figure 1.4: Hands creating gesture form extending between the fingers and in the length of the centrifugal motion. From Kendon (2004:192). Courtesy of Adam Kendon and Cambridge University Press. Copyright © 2004, Cambridge University Press.

Figure 2.1: (Part of) gesture form and articulator form (simplified).

Figure 2.2: (a) Acting opening a window (reproduced from Müller 2010:162); (b) representing how a car hits the wall (ibid., p. 161); (c) moulding a round bench (ibid.); (d) drawing the contour of a picture frame (ibid.). Drawings by Mathias Roloff (www.mathiasroloff.de). Courtesy of Cornelia Müller and De Gruyter; through Copyright Clearance Center. Copyright © 2010, Walter De Gruyter GmbH.

Figure 2.3: Two wiggling fingers representing walking legs occur in gesture as well as in sign language. From Liddell (2003:305; excerpt). Courtesy of Cambridge University Press. Copyright © 2003, Cambridge University Press.

Figure 2.4: Continuum of the degree of similarity between hand and referent within gestures that map in three dimensions.

Figure 2.5: Chaplin’s body curved adjacent to a gearwheel (Modern Times, Chaplin 1936; screenshot JH).
Figure 2.6: Iconicity and conventionality in signs (redrawn based on Mandel 1977:60; axis direction changed). ........................................................................................................................................... 32
Figure 2.7: Holding an imaginary object. From McNeill (1992:148, fig. 6.1.1 (in original), see also McNeill 1992:14, fig. 3.1). Courtesy of David McNeill and University of Chicago Press. Copyright © 1992, University of Chicago Press. ........................................................................................................................................... 34
Figure 3.1: A vector is part of the conceptualisation of pointing. From Pfeiffer (2010:17). Courtesy of Thies Pfeiffer. ........................................................................................................................................... 41
Figure 3.2: Kinds of simple pointing gestures. ........................................................................................................... 43
Figure 3.3(a-d): Trace leaving of different shapes in different directions. Lines (a,b) and surfaces (c,d) in parallel (a,c) and non-parallel (b,d) motion. ........................................................................................................................................... 46
Figure 3.4: Static (a) vs. dynamic (b) representation of a circle. From Calbris (1990:51; excerpt). Courtesy of Indiana University Press. Copyright © 1990, Indiana University Press. ........................................................................................................................................... 47
Figure 3.5: Drawing. See Figure 8.24 for a Gesture Form Analysis of this example. From Goodwin (The Body in Action (2003b:21), Palgrave Macmillan; excerpt). Reproduced with permission of Charles Goodwin and Palgrave Macmillan. This material may not be copied or reproduced without permission from Palgrave Macmillan. ........................................................................................................................................... 49
Figure 3.6: Drawing. See Figure 8.24 for a Gesture Form Analysis of this example. From Goodwin (The Body in Action (2003b:21), Palgrave Macmillan; excerpt). Reproduced with permission of Charles Goodwin and Palgrave Macmillan. This material may not be copied or reproduced without permission from Palgrave Macmillan. ........................................................................................................................................... 49
Figure 3.7: Surface-contact gestures. ........................................................................................................................................... 51
Figure 4.1: The 2½D sketch. From Marr & Poggio (1979:307). Courtesy of The Royal Society; through Copyright Clearance Center. Copyright © 1979, The Royal Society. ........................................................................................................................................... 67
Figure 5.1: Dimensions of the ground in a figure-ground construction. On the basis of Talmy (2000:391).75
Figure 5.2: Forms of one and two dimensions in topological distinction. ........................................................................................................................................... 76
Figure 5.3: Profiled hand (bold oval) within immediate scope within maximal scope. From Langacker (2008:64; excerpt). Courtesy of Oxford University Press. Copyright © 2008, Oxford University Press ........................................................................................................................................... 77
Figure 6.1: The relation between gesture and sign within the broader context of bodily action......... 78
Figure 6.2: (a) A |football team| sitting down represented by curled fingers being berated by (b) their |coach| represented by the index finger of the right hand. From Liddell (2003:310; excerpt). Courtesy of Cambridge University Press. Copyright © 2003, Cambridge University Press. ........................................................................................................................................... 80
Figure 7.1: Summary of preliminary structure of gesture types from Chapter 3. .............................................. 84
Figure 7.2: Simplified distinction of gesture types according to the profiled shape. ............................................ 85
Figure 7.3: Four different types of form in gestures conceptualisation. .............................................................. 87
Figure 7.4: One point (a) divides 1D space, two points (b) enclose 1D space. .................................................... 89
Figure 7.5: One straight line (a) divides 2D space, two (or more) straight lines (b) or one curved line (c) enclose 2D space. ................................................................................................................................................. 90

Figure 7.6: One plane (a) divides 3D space, two (or more) planes (b) or one curved surface (c) encloses 3D space. ................................................................................................................................................. 90

Figure 7.7: Forms of 0-3D affording division across one-to-three-dimensional space. The simplest forms (lowest number of dimensions) that afford enclosure are in bold face. .......................................................... 91

Figure 7.8: Prominent affordances of form dimensions in three-dimensional space. ......................... 92

Figure 7.9(a-d): Forms of different spatial dimensions in gestures. ......................................................... 93

Figure 7.10: Gesture Form Analysis coding scheme. ............................................................................... 97

Figure 7.11: Overview of dimensions of input and output form across static and dynamic gesture form operations. ................................................................................................................. 107

Figure 7.12(a,b): Stimulus object (ring) and ambiguous gesture either representing in 3D or enclosing the imagined object. ........................................................................................................ 108

Figure 7.13: Gesture form operations and approximatively corresponding semiotic notions. .............. 116

Figure 7.14: Gesture form operations grouped by three distinct kinds of spatial relations. ....................... 118

Figure 8.1: Objects of Study 1. .................................................................................................................... 120

Figure 8.2: Gesture types akin to pointing (blue frame) and drawing (green frame) in some functional gesture categorisations. ........................................................................................................ 122

Figure 8.3: Diagram showing spatial operations (3rd level of the hierarchy) as a result of specific forms (1st level) affording specific spatial structures (2nd level). ................................................. 123

Figure 8.4: Simplistic hierarchical distinction of gestures that direct, point and draw. .......................... 125

Figure 8.5: Structure of a tree diagram capturing the rising complexity of gesture types, in which each hierarchical level cumulatively adds one operation to the necessary spatial processing. 125

Figure 8.6(a,b): Representing in 1D indicating the angle of a drainage pipe. Adapted from Enfield (2009:7). Courtesy of Nicholas Enfield and Cambridge University Press. Copyright © 2009, Cambridge University Press. (In the visual part of Gesture Form Analysis, the original, un-annotated, picture is reproduced on the left (a), while the zoomed picture with the form annotations is located on the right (b). This is done to clearly separate the original material from the annotations from the Gesture Form Analysis.) ........................................................................ 126

Figure 8.7: Four gesture form operations, which are displayed with the same colours throughout all graphical Gesture Form Analyses of this dissertation. ......................................................... 127

Figure 8.8: Excerpt of a typology of pointing gestures in the form of a tree-diagram, including Directing as an immediate elaboration of Representing in 1D. (The following figures at the beginning of each section will build up on this diagram. The bold-faced dimensions below the name of the operation show the number of dimensions of the output form determined by that operation and the preceding ones. The operation of Articulator Profiling is not part of these diagrams since it occurs in all gestures analysed in this dissertation.) ................................................................. 127
Goodwin and Palgrave Macmillan. This material may not be copied or reproduced without permission from Palgrave Macmillan. ......................................................... 146

Figure 8.25: The resulting traces of Drawing gestures. From Mittelberg (2010a:368; excerpt). Courtesy of Irene Mittelberg and Equinox Publishing. Copyright © 2010, Equinox Publishing Ltd. .......... 147

Figure 8.26: Holding and Moulding in a simple form-based differentiation. ........................................... 149

Figure 8.27(a,b): Representing in 2D as a strategy to depict an oval shape (@nmLab). ......................... 150

Figure 8.28: Excerpt of a typology of Surface Contact Gestures (profiling the articulator as a 2D shape) in the form of a tree diagram. The following figures at the beginning of each section will elaborate on this. ................................................................................................................................ 151

Figure 8.29(a,b): Holding an unspecified imaginary object. ................................................................. 152

Figure 8.30(a,b): Holding by two opposing surfaces showing the |object|’s extent on one axis, performed by two flat hands and applied to the abstract category of Tweety and Sylvester cartoons. Adapted from McNeill (1992:148 fig. 6.1.1 (in original), see also McNeill 1992:14; fig. 3.1). Courtesy of David McNeill and University of Chicago Press. Copyright © 1992, University of Chicago Press. .......... 154

Figure 8.31(a,b): Holding a subcategory of a word class. Adapted from Mittelberg (2010b:131). Courtesy of Irene Mittelberg and Equinox Publishing. (Image also published in Mittelberg (2010a:353). Copyright ©2010, Equinox Publishing Ltd.) ........................................................................................ 154

Figure 8.32(a,b): Holding by two opposing surfaces showing an stimulus object’s thickness in vertical orientation performed by the index and thumb pad (@nmLab). .............................................................................. 155

Figure 8.33(a,b): Holding by two surfaces that fully enclose an |object| thus showing the |object|’s complete shape (@nmLab). ......................................................................................................................... 157

Figure 8.34: Holding by the complete inner surface of index and thumb, specifying an |object|’s 2D projection. ................................................................................................................................. 158

Figure 8.35(*1-*4): Different ways to indicate a circular form with different articulators (digit 1 is the thumb, 2 the index and 3 the middle finger). From Calbris (1990:51; excerpt). Courtesy of Columbia University Press. Copyright © 1990, Columbia University Press. ....................................................... 159

Figure 8.36(a,b): Pulling movement in a dynamic Holding gesture miming to open a door (@nmLab). 160

Figure 8.37: Typology of Surface Contact Gestures (profiling the articulator as a 2D shape) in form of a tree-diagram. ................................................................................................................................ ........... 161

Figure 8.38: Moulding a cylindrical form. Adapted from Sowa (2006:261; excerpt). Courtesy of Timo Sowa................................................................. 163

Figure 8.39(a-e): Depicting the Drop. Moulding (a,b) and a combination of Holding with the right and Moulding with the left hand (c-e; @nmLab). ......................................................................................... 164

Figure 8.40: Stimulus objects (a) Drop and (b) Torus. ........................................................................... 165

Figure 8.41: Moulding the Torus. The Adjacency is positioned last or second but last (white arrow; @nmLab). ......................................................................................................................... 165

Figure 8.42: Moulding a conic shape in 2D space (on vertical facing plane) (@nmLab). ......................... 167

Figure 8.43: Moulding a conic shape in 2D space (on mid-sagittal plane) (@nmLab)......................... 168

Figure 8.44: Moulding a conic shape in 3D space (on sagittal plane) (@nmLab). 169

Figure 8.45: Moulding a conic shape in 3D space (on vertical plane) (@nmLab). 169

Figure 8.46: Moulding a conic shape in 3D space (on horizontal plane) (@nmLab). 169

Figure 8.47: Moulding a conic shape in 3D space (on oblique plane) (@nmLab). 169
Figure 8.44: Moulding the flatness of an object’s surface (Study 2; @nmLab). ...........................................169
Figure 8.45: Two different gestures emphasising the gestural use of the head as representing a sphere in 3D (pilot of Study 2) (@nmLab). .................................................................173
Figure 8.46: The right hand representing the Egg in 3D (@nmLab). ...............................................................173
Figure 8.47: Different degrees of similarity when Representing in 3D in American Sign Language at examples of classifier handshapes reproduced from Schembri et al. (2005:275; excerpt). Courtesy of Oxford University Press; through Copyright Clearance Center. Copyright © 2005, Oxford University Press. .............................................................................................................173
Figure 8.48: (a) Broad Directing (triangle) and (b) Broad Pointing (triangle). ................................................177
Figure 8.49: (a) Broad Directing (cone) and (b) Broad Pointing (cone). .........................................................178
Figure 8.50: Broad Pointing (cone) directed towards the ground in front of the speaker. .........................179
Figure 8.51(a-c): Broad Pointing (“wide-hand point” in original). From Wilkins (2003:195; excerpt). 
Courtesy of David Wilkins and Taylor and Francis Group, through Copyright Clearance Center. 
Copyright © 2003, Taylor and Francis Group LLC Books. .................................................................179
Figure 8.52: Broad Directing (regardless whether triangle or cone) in flat, 2D space from a top perspective. ...........................................................................................................................................180
Figure 8.53(a,b): Surface Directing. From Wilkins (2003:197; excerpt). Courtesy of David Wilkins and 
Taylor and Francis Group, through Copyright Clearance Center. Copyright © 2003, Taylor and 
Francis Group LLC Books. .................................................................................................................182
Figure 8.54: Schematic illustration of Surface Directing. .............................................................................183
Figure 8.55(a,b): Surface Directing used for a path that includes climbing upstairs (@nmLab). ...........184
Figure 8.56: Examples of Surface Directing in route descriptions (a) from Wilkins (2003:197; excerpt) and 
(b) from Fricke (2007:110; 2013:741, 2014:1809; excerpt). (a) Courtesy of David Wilkins and Taylor 
and Francis Group, Copyright © 2003, Taylor and Francis Group LLC Books. (b) Courtesy of Ellen 
Fricke (www.ellenfricke.de) and De Gruyter. Copyright © 2003, Walter De Gruyter GmbH. 
Through Copyright Clearance Center. ....................................................................................................185
Figure 8.57: Applying the gesture form operation of Trace Leaving with a movement orthogonal to the 
profiled shape, resulting in different types of sketching, including Drawing. .......................................187
Figure 8.58: Possible interpretations of Surface Sketching extending away from the hand (c,d) or not 
(a,b), intersecting with a frontal plane (b, d) or not (a,c) ........................................................................191
Figure 8.59: Typology of gestures discussed in this dissertation based on Gesture Form Analysis; excluding various gesture types’ conceptual reduction of gesture space to one or two dimensions. ..................................................................................................................195

Figure 9.1: Study 1 demonstrated the partially competing maxims of a controlled vs. a naturalistic study design. ..............................................................................................................................................201
Figure 9.2: Age distribution boxplot. ............................................................................................................202
Figure 9.3: Capturing volume @nmLab (Photo: Matthias Priesters), see also Figure 9.4 for setup. ........203
Figure 9.4: Lab setup for Study 2 @nmLab including high-speed video cameras (HS), regular digital video cameras (DV), the participant (Part.), the interviewer (Int.) and the operator of the motion-capture system (Op.). ............................................................. 204
Figure 9.5: The six markers of the Minimal Marker Set. ................................................................. 206
Figure 9.6: Stimulus objects of Study 2. ....................................................................................... 207
Figure 9.7: Screenshot of Elan as shown to the participants. Bottom: tier with live annotations. .... 207
Figure 9.8: Illustration offering the participant to choose from three kinds of roundness: (a) circular, (b) cylindrical, (c) spherical. ........................................................................................................ 210
Figure 9.9: Hand Shape Coordinates for the right hand, oriented according to the 3-finger-rule. .... 212

Figure 10.1: Measure and Shape Gesture. From Sowa (2006:199; excerpt), where both gestures are called “connective hand-shapes”. Courtesy of Timo Sowa. ................................................................. 217
Figure 10.2: Measure Gesture. One specific group of gestures within the gesture type Holding........ 218
Figure 10.3(a-f): Possible referent shapes of a Measure Gesture. .................................................. 219
Figure 10.4: Shape Gesture, one group of gestures within the gesture type Holding. ..................... 220
Figure 10.5(a-f): Possible referent shapes of a Shape Gesture. ....................................................... 220
Figure 10.6: Curvature information deciding over participants perceiving Measure or Shape Gesture... 222
Figure 10.7: Screenshots of the expressive phase of 14 Measure Gestures (Study 1). ....................... 224
Figure 10.8: Screenshots of three Shape Gestures (Study 1). The data of this participant (taken to elaborate Study 2) did not include more Shape Gestures (for more examples of Measure and Shape Gestures, see Appendix 7). ......................................................... 224
Figure 10.9: The variable Curve inversely proportional to the distance spanned by the finger............ 226
Figure 10.10: Combinations of pinkie, ring and middle finger occluding the inner contour of index finger and thumb .......................................................... 227
Figure 10.11: Pinkie, ring and middle finger being raised, displaying the inner contour of index finger and thumb .......................................................... 227
Figure 10.12: Pinkie Curl as the angle from a side-perspective ....................................................... 227
Figure 10.13: Same Curl value for inclining a straight finger (preceding figure) as for curving the finger (this figure) .............................................................................................................. 228
Figure 10.14: List of quantitative exclusion criteria for the Measure and Shape Gesture annotations. The order in this list is also the order of execution. ................................................................. 232
Figure 10.15: Cohen’s and Fleiss’ $\kappa$ respectively for all combinations among the individuals interpreting Measure and Shape Gestures .......................................................... 232
Figure 10.16: Significant differences in means of Measure and Shape Gestures .............................. 233
Figure 10.17: Boxplot distributions of Measure Gestures (“no” curvature) and Shape Gestures (“yes” curvature): (a) index Curve; (b) pinkie Curl ................................................................. 234
Figure 10.18: (a) Scatterplot matrices of index Curve (distCm.Ah_pi_i) and pinkie Curl (angleDeg.Ah_pp_p), (b) superimposed by lines indicating the quadrants and by (c) example gestures of three configurations. .......................................................... 235

Figure 10.19: Frequency of Shape Gesture interpretations in 10 stages of decreasing pinkie Curl (low x-values more curled, high x values less curled). .......................................................... 236

Figure 11.1: Coarse analogy of different layers of analysis in speech and gesture. ........................................ 242
## List of abbreviations and coding conventions

### General

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>[object]</td>
<td>imaginary object</td>
</tr>
<tr>
<td>@nmLab, nmLab</td>
<td>(recorded at the) Natural Media Lab, RWTH Aachen, led by Irene Mittelberg</td>
</tr>
<tr>
<td>a.c.n.</td>
<td>ante christum natum, before Christ</td>
</tr>
<tr>
<td>af</td>
<td>articulator form</td>
</tr>
<tr>
<td>ASL</td>
<td>American Sign Language</td>
</tr>
<tr>
<td>b</td>
<td>body</td>
</tr>
<tr>
<td>cm</td>
<td>centimetres</td>
</tr>
<tr>
<td>D (oD, 1D, 2D, 3D)</td>
<td>dimension(s)</td>
</tr>
<tr>
<td>DEx</td>
<td>Data Extractor script</td>
</tr>
<tr>
<td>e.g.</td>
<td>exempli gratia, for example</td>
</tr>
<tr>
<td>EPAA</td>
<td>Elan Plugin for Automatic Annotation</td>
</tr>
<tr>
<td>etc.</td>
<td>et cetera, and so on</td>
</tr>
<tr>
<td>et al.</td>
<td>et alii, and others</td>
</tr>
<tr>
<td>et seqq.</td>
<td>et sequantia, and those that follow</td>
</tr>
<tr>
<td>gf</td>
<td>gesture form</td>
</tr>
<tr>
<td>i.a.</td>
<td>inter alia, among others</td>
</tr>
<tr>
<td>i.e.</td>
<td>id est, that is</td>
</tr>
<tr>
<td>ibid.</td>
<td>ibidem, in the same place</td>
</tr>
<tr>
<td>JH</td>
<td>Julius Hassemer, author</td>
</tr>
<tr>
<td>MoJa</td>
<td>Motion-capture-analysing Java script</td>
</tr>
<tr>
<td>p.</td>
<td>page</td>
</tr>
<tr>
<td>sec</td>
<td>second</td>
</tr>
</tbody>
</table>
### Gesture Form Analysis

<table>
<thead>
<tr>
<th>Code</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>AD</td>
<td>Adjacency (general)</td>
</tr>
<tr>
<td>AP</td>
<td>Articulator Profiling</td>
</tr>
<tr>
<td>CN</td>
<td>Continuation</td>
</tr>
<tr>
<td>DI</td>
<td>Adjacency: Division</td>
</tr>
<tr>
<td>EN</td>
<td>Adjacency: Enclosure</td>
</tr>
<tr>
<td>IN</td>
<td>Intersection</td>
</tr>
<tr>
<td>SP</td>
<td>Shape Profiling</td>
</tr>
<tr>
<td>TL</td>
<td>Trace Leaving</td>
</tr>
</tbody>
</table>

### Operator symbols

- $\in$: is element of
- $\land$: logical 'and'
- $\lor$: logical 'or'
- $\cap$: geometrical intersection, the space shared by two forms
The difference of dimensionality means that signages can be iconic to an extent to which languages cannot ... so that it is perhaps more revealing to put the difference the other way around, as a limitation of spoken languages. Indeed, the dimensionality of signing is that of life itself, and it would be stupid not to resort to picturing, pantomiming, or pointing whenever convenient. (Even when speaking we do this: for example, we utter a demonstrative such as there, which indicates relative distance but not direction, and supplement it by a pointing gesture that indicates direction but not distance.) But when a representation of some four-dimensional hunk of life has to be compressed into the single dimension of speech, most iconicity is necessarily squeezed out. In one-dimensional projection, an elephant is indistinguishable from a woodshed. Speech perforce is largely arbitrary; if we speakers take pride in that, it is because in 50,000 years or so of talking we have learned to make a virtue of necessity.

C. F. Hockett 1978:274-5
Introduction

This dissertation develops the notion of gesture form as being a necessary part of gesture conceptualisation. The question of “What is gesture form?” is a central element of the fundamental (in the sense of basic theoretical) research in the domain of gesture studies. Despite considering the subject matter from within the interdisciplinarity of gesture studies, the theoretical roots of this dissertation mainly stem from within the area of Cognitive Linguistics, and are applied specifically to the spatial modality of gesture.

The premises of a gesture-form perspective are gesture categorisations proposed by Müller (“modes of representation”; 1998; 2010; 2014) and Streeck (“practices”; 2009). More generally, academic theories on conceptual operations such as metaphor (i.a. Lakoff and Johnson 1980; 1999) and metonymy (i.a. Jakobson 1987/1956; Lodge 1977) and their application to gesture (metaphor: i.a. Cienki & Müller (Eds.) 2008; Forceville & Urios-Aparisi (Eds.) 2009; Müller & Cienki 2009; metonymy i.a. Mittelberg & Waugh 2009) guide the approach at hand. Other key influences are works focussing on pointing (i.a. Kita (ed.) 2003, specifically contributions by Wilkins and Kendon), a Gestaltist perspective on perception (e.g., Arnheim 2004/1969; Gibson 1977), the application of blending theory (Fauconnier & Turner 2000) to gesture and sign (i.a. Liddell 2003; Dudis 2007) and the systematic form analysis of “sketching” (Mandel 1977). Finally, Talmy’s (e.g., 2000) understanding of topological schemas supplies one structural pillar to the proposed Gesture Form Analysis, while parts of his current work (Talmy, forthcoming) overlap with operations of gesture form proposed in this dissertation.

Some gesture studies terminology

While specific terminology will be introduced in the respective sections, this section touches upon a few most basic definitions of gesture studies’ standard terminology.

"'Gesture', we have suggested, is a name for visible action when it is used as an utterance or as a part of an utterance." (Kendon 2004:7). Gestures are often performed while speaking, but can also stand for themselves. Body parts involved in gesturing are called articulators.

The temporal measure of a gesture unit (also gesture excursion) is divided into three main sequences, the gesture phases: preparation, expressive phase (Kita, van Gijn & van der Hulst 1998), and retraction. One gesture unit can have multiple expressive phases (often including respective preparations), and is limited, in the beginning and in the end, by the articulator being in rest position. When the expressive phase contains movement, it is called a stroke.
When the hand is held still, the expressive phase is called an independent hold. One expressive phase, including its corresponding preparation, is called a ‘gesture phrase’. Between any of the named phases, the articulator can be held in position for a period of time. This gesture phase is called a hold.

For example, the hand may be at rest hanging down at the side of the body (rest position). As soon as the hand (the articulator) is raised in order to perform a gesture, the gesture unit begins with the preparation. In the expressive phase that follows, “the form of the body movement is associated with the information to be conveyed” (Kita, van Gijn & van der Hulst 1998:28). In the stroke of the gesture in Figure 0.1, an index-finger pointing hand (index extended, other fingers curled in) is moved from left to right, drawing a line in the air: after this, the articulator remains in its position for a brief moment (hold), a new preparation brings (not illustrated in Figure 0.1) the hand into the starting position for the next expressive phase. In this expressive phase, the hand points: It is held for one second with the index finger extended (independent hold). Finally, the arm relaxes so that the hand swings downwards (retraction) and reclines at the side of the body (rest position), ending the gesture unit.

Figure 0.1: A gesture drawing a line (depicted in red) in the air.

Examples of gesture classifications

The most popular gesture classification is McNeill’s classification of iconics, metaphorics, beats, cohesives and deictics (McNeill 1992). Iconic gestures are similar to what they refer to—in aspects of shape, movement trajectory, or combinations thereof (e.g., moving the hand as if bending something being somewhat similar to the physical action of bending). Metaphorics are also similar to a referent; only that the referent is abstract (moulding a round shape in the air to express that a story was well-composed, ‘round’). Beats are movement impulses that mark a moment in time, an “expressive point” (Sowa 2006:110; e.g., emphasising a specific word in the accompanying speech). Cohesives are similar to beats in that they structure discourse and are variable in their form (e.g., marking a brief topical excursion by leaving a specific rest
position, performing a set of gestures and returning to the rest position when returning to the main topic). Deictics are pointing gestures (e.g., pointing at a book while saying “That’s mine!”).

McNeill’s gesture categorisation stands out by being used cross-disciplinarily (i.a., Liebal, Müller & Pika 2007; Poggi 2008; Sowa 2006), but it is also seen as problematic because it compounds functional, form-based and semantic criteria (Müller 1998; Cienki 2005). Also, McNeill’s different categories of gestures are not clear-cut types. Rather, a gesture belongs to more than one category at a time—to different extents (McNeill 1992; 2005). This is problematic for a quantitative approach that involves counting discrete quantities of gesture types.

Other classifications are more discipline-specific. Gesture studies is a domain in which scholars of areas as varied as Linguistics (especially Sign Linguistics and Cognitive Linguistics), Psychology, Ethnology, or Computer Science investigate “visible action as utterance” (Kendon 2004). Due to the interdisciplinary nature of the young and booming field of gesture studies, analyses often focus on aspects important within the background of the investigator’s original discipline. By way of example, Psychology-influenced classifications include or focus on self-adaptors like scratching one’s nose or fidgeting (i.a., Ekman 1969; Lausberg & Sloetjes 2009), since these can indicate stress or other psychological states, while other gesture research ignores these kinds of movements altogether (i.a., Krauss, Chen & Chalwa 1996; Schegloff 1984; Sacks & Schegloff 2002).

There are also more categorisations applicable to cross-disciplinary research. In order to objectify analyses, some approaches scrutinise verifiable, physical, form features of bodily articulation, above all in Sign Linguistics (e.g., Stokoe 2005/1960; Hanke 2004; Johnston & Schembri 2007). Similar efforts are also to be found in gesture studies (e.g., Bressem 2013; Gibbon et al. 2003; Martell 2005; Sadeghipour & Kopp 2012). Although these categorisations allow for comparative gestures analyses, no notation system is consistently applied across many of the disciplines that constitute gesture studies literature.

---

1 Although the study of gesture dates back at least to 100 a.c.n. (Quintilianus 1922/100-35 a.c.n.), ‘modern’ gesture studies developed only in the 20th century (for a historical synopsis consult Kendon 2004).
Besides the physical features of the body, there is another kind of form in gesture, the imaginary, or fictive\(^2\), form that an observer conceptualises\(^3\) when seeing a gesture (e.g., the red line in Figure 0.1). This form is the spatial information conveyed by a gestural movement, here called gesture form. Categorisations on the basis of gesture form exist (Müller 1998; 2010; 2014., Sowa 2006), but are not valid for all kinds of gestures (e.g., pointing gestures), among other aspects (see Section 2.2 and 3.4). Working towards such a categorisation of clear-cut gesture types is a main aim of this dissertation.

Several steps will be taken before proposing a gesture typology: first, defining gesture form in contrast to articulator form (Chapter 1). The second step is gathering form distinctions in the literature (Chapters 2-6), proposing operations that capture form distinctions and propose a scheme that can code gesture form operations and gesture form (Chapter 7, resulting in a definition of gesture form in Section 7.3), and successfully conducting the proposed Gesture Form Analysis on a collection of naturalistic data (Chapter 8). The last section of Chapter 8 summarises these analyses in a preliminary typology that captures all gesture types found in the data as well as in examples from the literature.

The three parts of this dissertation

Part I (Chapters 1-6) is a literature review of different disciplines with the leading question of which forms are being used systematically to describe different kinds of gesture, and how. Theories are examined in their mostly implicit foundations with regard to gesture form, which will be drawn upon in the theory of Gesture Form Analysis.

As a prerequisite for a literature review from different disciplines relevant to gesture studies, Chapter 1 introduces the two kinds of form whose distinction this dissertation rests upon: articulator form versus gesture form. Making this distinction is not only of theoretical interest, but also affects practical gesture analysis by introducing a new intermediate step. The chapter includes touching upon all potential senses of gesture perception and limiting the current approach to visual perception.

---

\(^2\) I use the term “fictive”, coined by Talmy for example in the notion of “fictive motion” (Talmy 2000:99-175), to stress that, at least in the mind of an observer, the motion ‘exists’. According to Talmy (ibid., p. 100), the “term ‘fictive’ has been adopted for its reference to the imaginal capacity of cognition, not to suggest (as perhaps the word ‘fictitious’ would) that a representation is somehow objectively unreal”.

\(^3\) The terms ‘conceptualise’ and ‘conceptualisation’ are used as in the linguistic work of Langacker: “In cognitive semantics, meaning is identified as the conceptualization associated with linguistic expressions” (Langacker 2009:4). Talmy (e.g., 2000; 2012; forthcoming) uses “conceptualisation” for language and including gesture.
Chapter 2 shows the roots of the proposed Gesture Form Analysis. The hands, and potentially all other articulators, are able to create fictive forms in the mind of an observer in different ways. In drawing a line in the air with an extended index finger, our hand performs a gesture that is analogous to using a finger to draw in the sand. When we mould an hour-glass shape with flexibly curved hands in the air, the hands perform a gesture of a different type and create form in another way. In trying to differentiate these and other gesture types, the use of so-called 'form parameters' (e.g., the hands' location, motion, configuration) is critically reviewed (Section 2.1). Work on various gesture types is shown to give insights into the variety of the gestural toolkit available to different articulators, especially the hands (2.2). Similar to Cognitive Linguistics’ notions of iconicity, metaphor and metonymy, there is a specific relation between articulator and referent in a gesture. Iconicity (2.3) and metonymy (2.4) are 'spatialisable'—that is operationalisable as spatial, geometrical relations—and thus applicable to Gesture Form Analysis. They describe necessary cognitive processes for understanding and discriminating gesture types. Yet, these and other semiotic relations lack the systematics with regard to the geometry in three-dimensional gesture space—unlike Mandel's matrix of "sketchings" (Mandel 1977; Section 2.5). As a first step in the direction of systematising semiotic notions for application in the spatial domain of gesture, the interaction of conventionality, metaphor and metonymy is illustrated in a single gesture that allows for inferring aspects of the shape of an imaginary object (2.6).

Chapter 3 introduces the current approach within the framework of embodiment (3.1). Three main sections examine the employment of the form descriptions of 'point', 'line' and 'surface' in analyses of pointing (3.2, 3.3), drawing (3.3), as well as holding and moulding gestures (3.4), and the way gestures can instantiate an entire three-dimensional form (3.5). Summarising the use of these geometric descriptors, their distribution does not appear to be equivocal (3.6), but following a systematics to be described later, in this dissertation’s theoretical contribution in Part II.

Among the different forms encountered in gesture analyses, Chapter 4 (of Part I) identifies two-dimensional surfaces to be central to the processing of sensory data of the senses relevant for gesture perception. This is equally true for visual perception, the sensory system most important in perceiving gestures, as well as for tactile perception, central to the actions that gestures imitate in the gesture types holding and moulding. Two-dimensional abstract forms as part of perception, are calling for a definition of the word ‘abstract’ (4.1) and motivating an excursion into the kind of spatial information that is processed in visual and tactile perception.
including how the specific form of the two-dimensional surface is crucial to a theory of affordances (Section 4.2).

The last two chapters in Part I briefly touch upon other important theoretical grounds. Chapter 5 reviews the systematic occurrence of forms revealed by how language structures spatial concepts. One- to three-dimensional forms are implied in the use of specific prepositions, potentially to reduce cognitive load when only specific aspects of spatial extension are relevant in an utterance (5.1). Moreover, the ability of language to focus on one structural unit within a complex entity (profiling) is a candidate for load-reduction that can be adapted to the spatial modality of gesture (5.2).

Chapter 6 refers to work in branches of Sign Linguistics for accounts on conceptual relations between action, gesture and sign (6.2). Conceptual Integration Theory is applied in order to map mental concepts onto the signing articulators and other physical objects, for example in classifier constructions. From these approaches, the terms real space and imaginary or mental space are adopted (6.3).

Part II (Chapters 7 and 8) follows up on the discussed theoretical aspects in an outline of the theory of Gesture Form Analysis. Chapter 7 presents the two principles of gesture form, its primitives and its operations. Chapter 8 applies Gesture Form Analysis to data of the dissertation study and gesture studies literature.

Part III (Chapters 9-12) reports on the methodology of Study 2 (Chapter 9) and its results, testing the central assumption of Gesture Form Analysis: whether gesture form deviates from the bare articulator form as predicted by the proposed operations (Chapter 10). Chapter 11 discusses the results of Study 2 and Chapter 12 closes by collecting the loose ends of a model of gesture form.
Part I:

Aspects of form in gesture research, and the theoretical bases of gesture form
1. Locating gesture form

The following sections will introduce the notion of gesture form. Gesture form is the spatial information conveyed in a gesture, abstracted away from the complex information of the gestural articulator moving through space (articulator form). Gesture form is derived from articulator form, which is the reason to begin by first defining and characterising the notion of articulator form, in a phenomenological perspective.

As ‘form’, more generally, I understand any spatial information, from the mere location of a point up to complex and dynamic information such as the rotation of a surface or the distortion of a three-dimensional entity, as in a wringed sponge. ‘Shape’ and ‘form’ both stand for ‘spatial entity’, potentially moving in space. ‘Shape’ is used especially for rigid entities and specific states (snapshots) of the complex movement of a non-rigid entity. ‘Form’ is then used more generally, including abstract descriptions of motion and non-rigid entities (regardless of whether they change their configuration by curving, bending or other distortions). Otherwise, both terms are interchangeable.

1.1. Locating the research interest within the phenomenon of gesture

Objectively, gestural action can be captured by the position of body parts as a function of time. The position and motion of these body parts is what will be called articulator form. Since body parts are physical objects, articulator form can be described in high detail or captured by motion-capture technology. The physical phenomenon articulator form, treated in this section, will be the basis to infer gesture form (Section 1.5; Chapter 7 and 8), which is this dissertation’s core topic.

Articulator form is defined by how specific articulators are moved or held still in gesticulation. This dissertation limits itself to the visual perception of ‘rough movement’. Thus, the otherwise general claims about the perception of gesture exclude other kinds of perception, and ‘fine movement’. To precisely locate the research interest of this dissertation, this section will touch upon different senses in gesture perception and different granularities of movement.

As illustrated in Figure 1.1, two granularities of movement are distinguished: rough movement concerns motion aspects that are sufficient to describe most gestures and their communicative function. Examples of rough movement are bringing the index finger into a position in which it points to a referent, or moving the hand in a circle to depict the shape of a round object, or holding an open flat hand, palm up, to present a discursive element—or any
other kind of gesture description on this level of form detail. These kinds of movement description do not touch upon movement below the range of centimetres and fluctuations in the speed of a movement. Rough movement is the level of granularity of the physical phenomenon gesture that underlies most gesture analysis including in this dissertation.

Aspects of fine movement play a more subtle role. The speed pattern within a gesture is usually characterised by an acceleration of the hand during the preparation (Priesters 2012) and either an acceleration during the stroke to mark a “wipe gesture” (Neff et al. 2008:9), or a pause in acceleration marking a pointing gesture (Latoschik & Wachsmuth 1998; see also Kita, van Gijn & van der Hulst 1998). It is not only different gesture phases that are characterised by fine movement. The work of Laban (e.g., Laban & Lawrence 1974; Bartennieff & Lewis 1980) deals with aspects of fine movement that determine, for example, how effortful a movement appears to observers. Movements with “bound flow”, accelerating and decelerating quickly, stand for more effort and “reveal the readiness of the moving person to stop at any moment” (ibid., p.15).

Although fine movement is crucial in segmenting gesture phases and indicating movement qualities, it is not relevant in determining the shapes, movement trajectories, locations or other spatial information gestures are known to give information on. To remain within a realistic frame of investigation, fine movement is excluded since it does not seem to impact articulator form, nor its derivative, gesture form.

<table>
<thead>
<tr>
<th>Sense</th>
<th>Phenomenon</th>
</tr>
</thead>
<tbody>
<tr>
<td>Vision</td>
<td>Rough movement (articulator form): Approximate position of the articulator as a function of time</td>
</tr>
<tr>
<td></td>
<td>Fine movement: Speed change (acceleration, steadiness, abruptness of stops, shaking)</td>
</tr>
<tr>
<td>Hearing⁴</td>
<td>Auditory signals, e.g., by tapping, snapping, clapping, knocking.</td>
</tr>
<tr>
<td>Touch⁴</td>
<td>Force, temperature, vibration, e.g., touching the interlocutor with the hand.</td>
</tr>
</tbody>
</table>

Figure 1.1: Three senses relevant in gesture perception.

⁴ Since the investigation will not follow up on hearing and touch as parts of gesture perception, they are not differentiated into fine and rough.
In communication by gesture, articulators moving through space are perceived largely visually. This dissertation’s focus on visual perception is not to ignore other senses that are relevant for perceiving gestures (Figure 1.1). At least hearing and touch can contribute to gesture perception. For example, clapping hands or touching the interlocutor clearly involves audio- and tactio-ception. However, only few gestures (also) appeal to these senses. In this dissertation, the proposed theory and the empirical investigation will focus on vision as the central sense of gesture perception because the senses of hearing and touch are not analysable with the theory proposed here. Proposing a theory that includes all senses is a challenging task that, by far, exceeds the limits of this dissertation.

Imagining tactile perception will play a role, but not in the immediate sense of gesture perception. Rather, interpreting the body to be in interaction with invisible entities, a central part of gesture conceptualization, involves imagining tactile contact with an object and imagining body movements (e.g., Müller 1998; Mittelberg 2010b), which may include cognitive processes related to proprioceptory and other somatosensory sensations.

The approach at hand simplifies in reducing perception to the five traditional senses. For gustatory and olfactory perception, I am not aware of studies justifying their consideration in studying the perception of gesture. The scope of this dissertation does not allow for including other potentially relevant aspects of perception, for example those perceptory systems processing signals of nociceptors or mechanoreceptors.

The highlight of “articulator form” in Figure 1.1 is to identify the analysis of rough movement as the starting point of this dissertation, differentiating it from other senses of gesture perception and contrasting it to aspects of fine movement. Scholars, who regard the form of a gesture as the basis of analysis—some explicitly using the term “form-based” (e.g., Holle, Obleser & Gunter 2011; Kopp, Sowa & Wachsmuth 2004; Ladewig 2011; Müller 2010; 2014; Müller, Ladewig & Bressem 2013; Sowa 2006)—use the word “form” to refer to the physical shape and movement of the articulators. “Gesture form”, then, is also not a fix technical term. Rather, “gesture” and “form” are freely combined, for example in “gesture’s form” (Bressem 2013), “form of the gesture” (Gerwing & Bavelas 2004), “gestural form” (Kendon 2004), or “gesture form” (McNeill 2005), i.a.), sometimes also within one text.

Some scholars specifically use the term “gesture form” to stand for the articulators’ physical shape and movement (e.g., Duncan 2002; Parrill 2008), but I do not want to limit gesture form to the physical form. A description of physical movement is not gesture-specific. Any action can be described by configuration and movement of body parts, not only gestures. Reserving a
gesture-specific term for those forms that are physically present would thus not be ideal. Rather, the form of a line in someone drawing a line in the air is gesture-specific; indeed, the imaginary line is the essence of what is being communicated by this gesture. Such imaginary forms should be included in gesture form. In some gesture types (miming someone’s actions that do not involve an object), gesture form may be limited to the physical form of the articulators, but in many others (e.g., drawing a line in the air; see Chapter 8 for a variety of examples), non-physically present imaginary forms are actually foregrounded. The rest of this dissertation is an attempt to offer a framework for analysing gesture form as consisting of both physical and imaginary forms.

‘Articulator form’ will be limited to the physical movement of articulators. The term ‘gesture form’ will be used for all the forms that are part of interpreting a body movement as an utterance or as part of it (Kendon 2004). The theoretical starting point of the investigation of different kinds of form in gestures is the non-equivalence of the physical form of any movement, articulator form, and all the forms that are part of the cognitive processing of movements interpreted to be a gesture: gesture form. The differentiation between articulator form and gesture form is a theoretical decision that impacts the basic structure of gesture analysis as shown in the rest of this chapter. Some approaches implicitly treat gesture form (though not using this term) as a separate analytic process, for example when drawings of gestures are illustrated with lines, abstracted away from the shape of the articulator (i.a., Bressem 2013; Efron 1941; Mittelberg 2010a; Müller 2008b) or when, for different gesture types, specific areas on the hand are marked as “meaning-baring” (Sowa 2006:91) and insights of such approaches will be used for the theory of Gesture Form Analysis.

In analogy to spoken language, articulator form would be the acoustic signal (physical phenomenon exposed to human perception) and gesture form would be the phonemic material (the meaning-differentiating part of a morpheme). This analogy will be touched upon in the discussion (Section 10.5), but is not the basis for the modality-specific perspective on gesture followed in this dissertation.

---

5 In other words, someone failing to see the form of a line in such a movement actually fails to see this movement as the gesture it is.

6 A gesture type (see Chapter 2)—being composed of gesture form (consisting of constituent forms, see Part II)—would then be analogous to a morpheme.
1.2. Two levels of gesture analysis

In this and the following section, the structure of existing gesture analyses will be examined with regard to the different kinds of form that are part of it.\(^7\) In general, a two-level analysis underlies detailed qualitative approaches as well as approaches that merely quantify gestures with the help of predefined categories. The next section will propose an additional level of analysis—not for a specific branch of gesture studies, but for the analysis of gesture in general. To work towards explicating and implementing this additional level in existing analytic methodologies is the aim of this dissertation.

In gesture studies, the phenomenon of gesture is usually described on two levels—at least I am not aware of an explicit use of a third level. In consonance with the approach in this dissertation, it is widely assumed that the physical phenomenon is the basic, objective level of analysis. A more detailed description, closest possible to the physical reality, is associated with a more objective analysis (Bressem 2013). Not only human analysis can scrutinise the movements that make up a gesture, but also computer-aided methods such as motion-capture devices of different kinds. Whether including or excluding the help of automatic computation, whether more or less detailed, the first level of analysis deals with the movement of the articulators in space, the “anatomic and visuo-spatial characteristics” (Grandhi, Joue & Mittelberg 2012:1): it deals with the physical articulator form.

How the hands or other articulators are being moved is generally a basis for inferring gesture meaning. Depending on the background and research interest, this interpretation may include speech and various other types of context as well as psychological, ethnographic or other factors relating to specific language communities that influence the possible meanings of a gesture. Analyses of form are sometimes limited to the articulator form—sometimes in great detail, sometimes only with regard to a single aspect. Other analyses specify imaginary forms that can be inferred from gestural action, for example in pointing gestures (\textit{i.e.}, Fricke 2007; Kendon 2004; Sowa 2006), drawing gestures (\textit{i.e.}, Bressem 2013; Efron 1941; Mandel 1977; Mittelberg 2004a; Müller 2008b; Streeck 2008), or gestures that hold or interact with imaginary objects (\textit{i.e.}, Andrén 2008; Grandhi \textit{et al.} 2012; Kendon 2004; Müller 2004; McNeill 1992; Mittelberg & Waugh 2009; Sowa 2006; Streeck 2008). While Sowa (2006) suggests form calculations for many specific gestures, a single system of spatial relations that systematically

\(^7\) The reference to ‘existing’ analyses is a strong generalisation that is adopted here only with regard to the question of whether the forms that are part of gestures analysis are explicitly distinguished into more than one layer, or not.
accounts for how these inferred forms emerge in different gesture types, does not yet exist. The
distinction between articulator form and gesture form is the foundation of any approach that
includes all relevant forms in gesture conceptualisation (across gesture types) in a single
framework, and investigates how the imaginary forms relate to the objective physical form.

1.3. Three levels of gesture analysis, introducing ‘gesture form’

This dissertation proposes an additional, intermediate level of analysis: Gesture Form Analysis.
Articulator form is conceptually processed, resulting in gesture form. Meaning, then, is
processed on the basis of gesture form. A central proposition of this approach is that the
physical movement of the body is only in specific cases the most relevant form in meaning
making. Gestures tend to convey meaning with the help of imaginary forms that are inferred
from the body movement. These imaginary forms constitute ‘gesture form’ and they may have
a very different shape than our bodily articulators, thus enabling us to convey forms that look
very different from our bodily articulators (e.g., two hands ‘moulding’ a big sphere in front of
the gesturing body) or are very distant from our bodily articulators (e.g., pointing to a remote
object). Not original in this dissertation are many of the insights on how abstract forms can be
derived from the articulating body, and these insights will be referred to beginning in Chapter
3. Original in this dissertation is the explicit distinction between gesture form and articulator
form and a proposal of how these spatial inferences may be integrated in one framework.

To justify the usefulness of an additional level of analysis, this level should fulfil two
criteria: it should offer a new perspective on the phenomenon and it should streamline the
analytical process so that the already intricate analysis does not unnecessarily expand in
complexity and labour. How gesture form allows for a new perspective on gesture will be
treated at length in the rest of this dissertation. For the current chapter, an explanation of how
three instead of two levels of analysis can streamline instead of inflate the analytical process, is
in place.

At least two kinds of complexity make gesture analysis a challenge. First, an astute
description of how all relevant parts of the body move, is a demanding and time-consuming
task. This is due to the complexity within the articulator form. Even the action of a single
complex articulator, for example the hand, is difficult to describe. Complexity multiplies when
one takes into account that every body part and the whole body itself can gesture and that
there is an interaction between all these articulators. Figure 1.2 gives an impression of the
complexity of multiarticulatory data in a short time frame (Hassemer 2009). As each
articulator in the figure is coded by descriptions on one tier only, a more detailed and

Figure 1.2: 1.5 seconds of Elan coding of multiple articulators annotated in one tier each (Hassemer 2009:68).

Also, no matter how flawless and well-elaborated a verbal description may be, seeing the video of a gesture (let alone being co-present) will always enrich the appreciation of a communicative act. To thoroughly capture the phenomenon, ever more detailed analyses are required, which are ever more difficult to overview and inevitably depart from the spontaneous impression vital for the perception of gestural action.

In seeming opposition to these complexities stands the often abstract and pleasingly simplifying character of gesture. For example, a complex motoric interaction of upper and lower arm may cause the index finger to move straight ahead through space, thus drawing an imaginary straight line in the air. A straight line is a highly abstract, simple form, which may in turn stand for a rich concept such honesty (e.g., Cienki & Müller 2008). Gesture is a vital part of human communication, not despite but because it reduces complexities. “The gesture limits itself intelligently to emphasizing what matters” (Arnheim 2004/1969:177). Gesture Form Analysis tries to capture part of the limited and thus essential, abstract information of the spatial modality gesture. Gesture form is independent of specific articulators on the one hand, and independent of specific context on the other. It is not only the relevant—and, as will be argued, even necessary—spatial information that is abstracted away from the complexity of the articulator form, but it is at the same time the core information that is put into the big picture.
of complex context (Figure 1.3). As the intermediate level of analysis, gesture form is given a pivotal role in the analysis of gesture conceptualisation. On a more abstract level, Gesture Form Analysis deals with the overlap between what is essential in the movement of articulators and what is the gesture-specific contribution in the overall utterance, adding to or confirming information gathered from speech and spatial context.

As a continuation of a form-based perspective of analysis, the main contribution of this dissertation is a methodology that explicitly differentiates articulator form (existing physically) from gesture form (overlapping only in part with articulator form, abstracting away from articulator form and inferring additional forms, see Chapter 7). The core analytic claims are:

gesture form is not equivalent to articulator form. They are separable steps of analysis. Gesture form comprises the forms that are conceptually necessary for understanding gestures. With regard to the “formulator” in Levelt’s speech-production model, being the intermediate step between conceptualisation and articulation of speech, gesture form could be analogous to “lemmas” in the speech formulator (Levelt 1983) with the crucial differences—specifically speech and specifically production.

Several scholars conduct gesture analyses that imply aspects of an intermediate analytic step that deals with derived or inferred imaginary forms (see Chapter 2 and 3). Kendon’s analyses might show the most rigor in differentiating physical and conceptualised forms of a

---

8 Lemmas “are specified semantically, as well as in terms of syntactic category and functional syntactic properties” (Levelt 1983:46). This is analogous to gesture form containing semantic content as well as being adapted to the constraints of articulatory movement. Yet, lemmas are part of a lexicalisation process, and accessing fixed lexical items (as a basis for generating “necessary and sufficient instructions for the motor executive programs”, (ibid, p. 47)) conflicts with the generation of analogous, non-discrete, gestures. An integration of Gesture Form Analysis with other models that build on Levelt and include gesture production (de Ruiter 2000; Krauss, Chen & Gottesmann 2000), is desirable. For example, de Ruiter’s “gesture planner”, connecting the “conceptualiser” with “motor control” (2000:294-8) could be argued to contain templates of gesture types constrained by the principles proposed in Chapter 7. While the formalisation efforts in Chapter 7 are intended to make Gesture Form Analysis compatible with information-processing models, proposing a means of integration exceeds the limits of this dissertation.
gestural movement. In the following example, he implicitly adheres to three levels of gesture analysis, providing separately descriptions of articulator form and gesture form:

As he says “bar” he lifts up his two hands, held with palms oriented outwards, with all digits extended but held together, thumb opposed but extended, which are then moved horizontally away from one another [Figure 1.4; JH]. The effect of this gesture is to mould a wide horizontal elongated object, which has a certain depth. The speaker thus ’creates’ the bar ...

Kendon 2004:191

Figure 1.4: Hands creating gesture form extending between the fingers and in the length of the centrifugal motion. From Kendon (2004:192). Courtesy of Adam Kendon and Cambridge University Press. Copyright © 2004, Cambridge University Press.

The reference to Figure 1.4 in the quote above is deliberately put at the same place as Kendon’s figure reference (original: “Fig. 10.12”; here reproduced in Figure 1.4). I assume, that the reference to the figure occurs exactly at this point in the description because Kendon has just finished describing the hand shape and how the hands move (articulator form)—a description that overlaps with what Figure 1.4 shows—and will now go on describing the imaginary form that this movement evokes (gesture form): “The effect of this gesture is to mould a wide horizontal elongated object, which has a certain depth” (Kendon 2004:191).

Only after both these parts, Kendon relates the analysed form to the context of a specific scene, the description of a door, and describes this form to be mapped onto a specific object, the security ‘bar’ that is to be placed on a door (third level of analysis: context-dependent interpretation of meaning). This incisive style of analysis implies an intermediate level of analysis as do other authors’ analyses (Cienki; Kita; Mittelberg; Müller; Streeck).

1.4. Reasons for a gesture-specific methodology

The approach taken in this dissertation considers speech as context for gestures (Kendon 2004; Mittelberg & Waugh 2009; Sweetser & Sizemore 2006; Willems, Özyürek & Hagoort 2007). Speech context can pointedly alter a gesture’s interpretation. Consider the following aphorism:
'whoever points with the index finger at someone (in the sense of accusing someone of something), should be aware that three fingers of that hand point back at you' (gloss JH). In pointing with the index finger, the middle, ring finger and pinkie are often folded under, thus being directed in the opposite direction: they point back at the gesturer. In a situation in which one person points at another and a third party utters this proverb, the proverb becomes the verbal context that shifts the focus of this regular index-finger-pointing hand away from the index finger and to the other fingers, thus inversing the pointing direction. The gesturer points back at himself (the accuser becomes the accused).

On the other hand, we understand our interlocutor on the phone or over an intercom without seeing gestures (Cohen 1977), and, when someone points at an object, we do not necessarily need speech to understand the gesture. Although natural communication often involves a vivid interaction of both modalities, both gesture and speech can also occur independently of one another and still be intelligible. I regard speech as context for gesture as a possible link, but not a necessary link.

The link of gesture and space, in contrast is of a mandatory kind. Gesture Form Analysis relies on a model of gesture form that consists of basic geometrical relations in three-dimensional space. Gesture form comprises spatial entities that are imaginary (though sometimes coinciding or overlapping with aspects of articulator form) and which, I suggest, structure gesture conceptualisation. In contrast to speech, space is a constituting part of gesture. It is not context, nor an optional associate: gesture can function without speech, but gesture cannot function without space. Space is the condition of possibility (Bedingung der Möglichkeit, Kant 1787:59) of gesture—for both gesture production and gesture perception. Thus, the model underlying Gesture Form Analysis depends on geometric relations and affordances of specific shapes. I rely on the notion of “affordances” as coined by Gibson (e.g., 1977). While Gesture Form Analysis attempts to describe how articulator form is conceptualised as gesture form (left arrow in Figure 1.3), and how gesture form is put into the context of the overall utterance (right arrow in Figure 1.3), the structure of this analysis is speech-, articulator-, and context-independent.

---

9 As a reaction to left-wing protests after the assassination of the political activist Dutschke in 1968, the Chancellor Kiesinger strongly condemned the protestors. President Heinemann addressed these denunciations by saying: „Wer mit dem Zeigefinger allgemeiner Vorwürfe auf den oder die vermeintlichen Anstifter oder Drahtzieher zeigt, sollte bedenken, dass in der Hand mit dem ausgestreckten Zeigefinger zugleich drei andere Finger auf ihn selbst zurückweisen.“

10 Proposed in Chapter 7 (see e.g., equation (13) and Figure 7.10)
The structural independence of gesture form is neither in conflict with integrated analyses of gesture and speech, nor with the interdisciplinarity of gesture studies. While gesture studies profits greatly from the diverse methodologies of different disciplines, there is also a gesture-specific core that demands specialised theory and methodology. A relevant part of this gesture-specific core is the investigation of operations that capture gesture conceptualisation in the light of the specific affordances of this spatial modality. These operations will be proposed as structural elements of Gesture Form Analysis. They are intended as a contribution for any study of gesture that requires distinctions of gestural form, regardless the disciplinary background of the particular research question.
2. Gesture form in different gesture types

The notion of gesture form relies on the understanding that, for every gesture type, an articulator such as the hand is perceived as imitating a specific practice (e.g., Streeck 2009). Which practice this is must be negotiated anew for every gesture (see Chapter 8 for an overview of different gesture types). The way the articulators are moved reveals the practice that the hand is imitating. Gesture types differ in the spatial relation between the hand (or another articulator) and the referent (or other spatial information communicated in this gesture, e.g., a direction in a route description). The literature cited in this chapter will be reviewed specifically with regard to different authors’ propositions about the spatial relations of hand and referent.

Section 2.1 will introduce the use of form parameters in gesture studies and distinguish them from gesture form. The parameters, including the configuration, location and motion of the hands (Stokoe 1960), are means to describe bodily action in general. Section 2.2 introduces Müller’s modes of representation (1998; 2010; 2014) as the starting point. Further sections will introduce the semiotic terminology of iconicity (2.3) and internal and external metonymy (2.4), which are already being used in gesture studies to describe spatial relations between articulator and referent. A single one of these spatial relations alone does not fully characterise one gesture type, they are combined within a gesture type. They build up on each other like a piece of chalk plus the direction of motion determine the shape of the trace left on the blackboard (the tip of a piece of chalk leaves a trace in form of a line, a piece of chalk pressed flat on the board leaves a trace in form of a surface, but only when moved laterally)—an interaction of spatial constraints that occurs in various gesture types (2.5). Section 2.6 analyses an instance of manual action, in which iconicity, metonymy, conventionality and metaphor are all part of creating gesture form of a single gesture.

2.1. ‘Form parameters’ describe articulator form, not gesture form

In 1960, Stokoe introduced the aspects of “hand configuration, position or location, and movement” describing manual action in American Sign Language (ASL). Since then, the so-called ‘form parameters’ have become a basic descriptive tool in sign languages—often extended by parameters of orientation, handedness and others. Also gesture researchers have fruitfully adapted the form parameters. Among the approaches that regard the form of a gesture to be an important part of the analysis, a detailed description of the gesture phenomenon hardly seems possible without Stokoe’s parameters (Andrén 2010; Calbris 1990;
Fricke 2007). Further, some approaches follow these parameters as driving force of their methodology (i.a., Bressem 2013; Harrison 2009; McNeill 1992; Parrill 2008).

For the theoretical study of gesture form, the question arises whether these parameters are also principles of gesture form. The answer is no. Form parameters are a means to comparably describe the physical form of manual action by capturing specifically the bare movements without any subjective interpretation, thereby successfully minimising the subjective interpretation of form. The parameters are an intuitive and productive way to describe and code how the hands move (or are held still), but they do not concern the way bodily action is interpreted. In other words, they describe only the articulator form of movements that are part of gesture or part of sign language.

Moreover, the parameters pose a problem for the endeavour of proposing a typology, since they are non-exclusive categories (Bakeman & Gottmann 1997). Already the possibility to extend the parameters, for example by the parameter orientation, nurses doubts about their discreteness. Stokoe himself sees these parameters as non-exclusive. He wants “to avoid the suggestion of mutual exclusiveness these words have in their ordinary uses” (2005/1960:20). In describing articulator form, the parameters can be interchanged: flapping down the index finger might be coded as a change in the ‘configuration’ of the hand, but there’s nothing apart from convention to prevent it from being called a rotational ‘motion’ of the index finger (hinging on the metacarpophalangeal joint), or a change of ‘position’ of the tip of the finger from one place to another. If the parameters are all interchangeable, they can be broken down to a description of articulator form that replaces all three without loss. Articulator form cannot only be captured by the standard Stokoean form parameters but for example also by (far less intuitive) values of the joints that make up the articulator in the case of angular motion capture devices or by the absolute position of any part of the articulator through time in the case of optical motion capture devices. Precise information on where all the parts of an articulator are at any point in time can exhaustively account for the entire articulator form, including the information offered by any combination of form parameters.

A description of articulator form is not concerned with how bodily action is used as “part of an utterance” (Kendon 2004:7). When action is used as utterance, we perceive a form that is abstracted away from articulator form. When someone moves the hand in central gesture space from left to right in an index-finger pointing configuration—described here in terms of the form parameters (location, movement, hand configuration)—I claim that the articulator form, as described, is non-equivalent to the form most saliently conveyed in this gesture,
gesture form. Such a motion of the articulator can draw an imaginary line in the air (see Figure 0.1), in the perception of any observer including the gesturer. This specific line, possessing a certain length, a position and possibly a certain degree of curvature, is part of what is meant by gesture form. The difference as well as the relation of articulator form and gesture form is the subject of this dissertation. Gesture form is claimed to be non-equivalent to, but reliant on, articulator form.

![Figure 2.1: (Part of) gesture form and articulator form (simplified).](image)

While the hand and the extended index finger is necessary to infer the gesture form of a line being drawn (Figure 2.1), the created line is, firstly, by far a simpler form than the form of an articulator as complex as the hand. It is in this case, secondly, not even strictly part of articulator form. The line is the consequence of interpreting the hand to leave a trace that is adjacent to one end of its linear extension; the trace being equally as long as the performed movement. Figure 2.1 simplifies the relation between articulator form and gesture form that, even in this straight-forward gesture, is multi-layered (see Section 8.4.7). Gesture form itself is made up of multiple constituent forms created in multiple steps, for which analyses in the literature provide indications (Chapter 3), and which will be teased apart in detail later (Part II of this dissertation).

Introducing a new definition of gesture form is not a terminological specification only; gesture form is to provide a spatial specification of other authors’ work particularly in reference to the relation between the visible forms of the hand and the invisible, abstract information of gestural communication.

\[\text{Figure 2.1: (Part of) gesture form and articulator form (simplified).}\]

\[\text{Introducing a new definition of gesture form is not a terminological specification only;}\]

\[\text{gesture form is to provide a spatial specification of other authors’ work particularly in}\]

\[\text{reference to the relation between the visible forms of the hand and the invisible, abstract}\]

\[\text{information of gestural communication.}\]

\[\text{\^* Strictly speaking, gesture form includes articulator form. This figure highlights the difference between}\]

\[\text{the two: articulator form not including the line in the air.}\]
Concepts of the intangible and invisible are only learned through percepts of tangible and visible objects, whether finally expressed to the eye or to the ear, in terms of sight or of sound.

Mallery 1881: Powers of signs compared to speech

The assertion of gesture form is that certain invisible, imaginary forms are a necessary part of understanding gestures. Any one gesture type (e.g., drawing) can be performed in manifold ways, but it is always characterised by the kind of imaginary forms that it creates (e.g., leaving a line as its trace). Thus, gesture form can help define gesture types (see the next two sections and Chapter 8).

2.2. Gesture types: the hands imitating different practices

A range of literature in gesture studies advocates the idea that the hand can create form in different ways; as if the hands had the possibility of becoming different tools (Andrén 2010; Mittelberg 2010a; b; Müller 1998; 2010; 2014; Streeck 2008; 2009).

This view entails a specification of the term ‘iconic’, which will be laid out in this section by defining gesture types that are iconic in different ways and, in the next section, by proposing different degrees of iconicity. Calling a gesture iconic is not very specific with regard to gesture form: the fact that some aspect of the gesture resembles some aspect of a referent makes for a very broad class of gestures. “The view that ‘iconic’ gestures uniformly function by way of some resemblance between signifier and signified is rejected” (Streeck 2008:285). Rather, there is “a heterogeneous set of practices” that all map iconically—but in different ways (ibid.; see also McNeill 2005:8). Clearly, many scholars differentiate within different kinds of iconicity and different kinds of iconic gestures (Andrén 2010; Fricke 2007; Mittelberg & Waugh 2009).

Building on their work, I will argue that there are different spatial relations within iconic gestures whose systematic description can yield insights on the structure of gesture conceptualisation. Note that neither Streeck’s nor the approach at hand reject the notion of iconicity in gesture, they rather attempt to specify the term ‘iconic’ across gesture types.

In a Vygotskyan tradition, Streeck emphasises that concepts in general, and language and gesture in specific, are rooted in “social practice” (Streeck 2009:30). This dissertation also adheres to this ‘praxeological’ view on gestures, so that also the abstract, formalised claims on form conceptualisation in gesture have their basis in the structure of human (inter)action (see Chapter 4). The spatial relations that are proposed for the processing of gesture come from an investigation of how for example the hands physically act on objects. The focus on form draws the attention specifically to “the complexity of the material structure” (Hutchins 2005:1555) of
the hand and the object they usually interact with. Further, the comparably free design of the empirical study is an effort to let the participants choose whatever kind of gestural means they like, not prompting to execute a specific gesture.

The notion of gesture form as promoted here, rests on Müller’s linguistic and form-based categorisation of four gestural modes of representation (1998:109-20) categorising gestures as “acting, molding, drawing, and representing” (Müller 2014:1691). Müller’s categories differentiate gestural function and also characterise as what kind of tool the hand is to be imagined (resorting to Gombrich, e.g., 2002/1960 and Mauss 1973). Seeing the bodily articulators as the human’s most basic instrument roots in Mauss’ “Techniques of the body” (1973):

I made ... the fundamental mistake of thinking that there is technique only when there is an instrument. ... The body is man’s first and most natural instrument. ... Before instrumental techniques there is the ensemble of techniques of the body.

Mauss 1973:75-6

Figure 2.2 shows examples of Müller’s four modes of representation (Müller 2010:160-2). In the mode of acting, for example miming how to open a specific type of window (reproduced in Figure 2.2a), the hand re-enacts the manual activity of opening a window. In representing, as in depicting how a car hits a wall by moving the fist against a flat hand (Figure 2.2b), the hands represent the car and the wall, showing the impingement of the two. In moulding, as in circumscribing an imaginary bench with slightly cupped hands around one’s body (Figure 2.2c), the hands enclose an imaginary referent, indicating its curved proportions by the space around which they curve in a circular movement. The hands “mold or shape a transient sculpture” (Müller 2014:1691; see also Section 2.4 and Mittelberg & Waugh 2009). In drawing, as in outlining the round contour of a picture frame with the tip of the finger (Figure 2.2d), the thus-created curved path characterises the frame as round.12

12 In recent work, Müller (2010; 2014) reduces the four modes to two fundamental modes of acting (including moulding and drawing) and representing, since “in the acting mode, the hands mime themselves, while in the representing mode they mime other entities” (Müller 2014:1696). This differentiation is not assumed as a basic distinction in this dissertation’s typology whose criteria are strictly based on the four dimensions of form (0-3D; see Section 8.5). An additional differentiation within the category of gestures that map three-dimensionally is possible. Either a binary one, as suggested for example by Müller, or a scalar one as proposed by Mandel and in Chapter 2.3 (see Figure 2.4).
Within Müller’s four modes of representation, I would like to draw attention to the different kinds of form that map from the hand onto a referent, or, more generally, the relationship between the physical form of the articulator (the creating form), and gesture form (the form being created by the articulator).

In acting and representing, the hand stands for the referent in its complete form. In miming how to open a certain type of window (reproduced in Figure 2.2a), the hand stands for a hand by instantiating what kind of movement needs to be performed (by a hand). In representing, as in showing how a car hits a wall, the hands represent aspects of motion (moving car, static wall, impingement of the two) and shape (bulky vehicle, upright wall). While the hands in representing and acting stand for distinct kinds of referents (Müller 2014:1691-4), they both map in their solid form. In representing, the hands can stand for any referent that they resemble in any aspect of shape or motion as “a kind of manual ‘sculpture’” (ibid., p. 1691), in acting, the hands in their dynamic form (including movement) stand for hands—performing a specific action. With regard to form dimensions, the two types are equal in that, in both, the hand represents in its solid, three-dimensional, form.

This is different in moulding and drawing (two categories that will be largely adopted in this dissertation’s typology). In moulding, as in circumscribing an imaginary round bench (Figure 2.2c), the hands enclose the imaginary referent, indicating its size by the space between the inner sides of the palms and conveying a curved shape by a circular movement. The space occupied by the hands is not the referent itself (as in acting and representing), the hands are external to the referent. An observer of this gesture type can infer the shape of the referent by imagining the hands sliding over its surfaces. In drawing, as in drawing the contour
of a picture frame (Figure 2.2d), even a smaller part of the hand is in contact with the imagined object. In other words, a smaller part of the hand creates gesture form. Only the tip of the finger traces a form that can then be mapped onto an object’s shape. In summary, we see a solid or three-dimensional overlap of hand and referent in acting and representing, whereas hand and imaginary object are in contact only by a surface in moulding, or by a point in drawing.

Müller’s gesture types (“modes of representation”; Müller 1998; 2010; 2014) overlap partially with Streeck’s twelve “depiction practices” (Streeck 2008). Both scholars see “drawing” and “acting” each as one type. Streeck’s “making” and “handling” are comparable to Müller’s “moulding”, and Streeck’s “pantomime” resembles Müller’s “acting”—with regard to the spatial relations relevant in this review. This partial correspondence calls for closer investigating gesture form in these types: What is the geometrical relation between the hands and the (imaginary) objects in these gesture types?

The examples of cited gesture types discussed above are structured by a very specific relation of the articulator and the imaginary object. These relations can be spatially described: the hand either stands for the referent in its solid shape or indicates aspects of the object’s shape by contact. While these relations vary from gesture type to gesture type, they seem to be the same within one gesture type. The kind of contact is equal, whether a curved or straight line is being drawn. Whether the shape being created is a square or circle, the description of a bee’s flight, or the drawing of a name, the contact between the articulator and the imagined object is point-like. This is articulator independent. It holds for drawing with the foot, a laser pointer or any other way of gestural drawing (Clark 2003:251-2; 1996).

Müller’s approach parallels the work of gesture researchers, who analyse the structure of how the form of (manual) action is interpreted (e.g., Andrén 2010; Calbris 2011; Fricke 2007; Goodwin 2003b; Hutchins 2006; Streeck 2009). The detailed analysis of the hand as different tools is called micro-analysis of different gestures (Müller 1998:114-27). Since the analyses of this dissertation will attempt to break the analysis of gesture types down to constituent elements, it could be called a nano-analysis. In the interest of not inflating the vocabulary of granularity, the approach here is termed Gesture Form Analysis. Still, part of the argument for a new theory is the wish to increase the analytic granularity with regard to form. Subscribing to Streeck’s account of heterogeneity of iconicity (above), this dissertation attempts to specify different kinds of iconicity with regard to the constituting spatial relations (Part II).
2.3. Iconic gestures can be more or less similar to their referent

The relation of articulator form and an imagined object in an iconic gesture can be characterised for example with regard to the degree of similarity: although acting and representing could be put into one form category since they rely on solid (three-dimensional) mappings, they are different in the degree of similarity between the articulator and the referent. The similarity is high when one mimes an action because the hand stands for a hand (Figure 2.2a). The case is different, and similarity progressively more limited, when my hand does not instantiate itself but another adult’s hand (still very high degree of iconicity), a baby’s hand, an ape’s hand, an animal’s claw, a machine’s gripper or just any object that has certain similarities with the articulating hand. As examples for a low degree of iconicity, the shape of a gun can be instantiated by the index finger and thumb extended, other fingers curled in, or the wiggling index and middle finger can stand for walking legs (as in the ASL example reproduced in Figure 2.3 from Liddell (2003))—both examples for Müller’s representing.

Figure 2.3: Two wiggling fingers representing walking legs occur in gesture as well as in sign language. From Liddell (2003:305; excerpt). Courtesy of Cambridge University Press. Copyright © 2003, Cambridge University Press.

The difference in similarity between acting and representing spans a continuum from the highest degree of similarity (hand stands for hand) to the least degree of similarity that still affords a mapping (fingers stand for legs). On this continuum, instances of acting and representing fall into different ranges: acting requiring more similarity and representing requiring less (Figure 2.4).

---

13 Unless stated otherwise, sign language examples are in ASL.
Figure 2.4: Continuum of the degree of similarity between hand and referent within gestures that map in three dimensions.\textsuperscript{14}

2.4. Two different kinds of metonymy in gestures

Apart from similarity and its different degrees, also the cognitive process of metonymy can be used to systematically tease apart spatial relations in gesture conceptualisation. Two kinds of metonymy have to be distinguished that are both part of the conceptualisation of gestures: internal and external metonymy. In the differentiation of these two types of semiotic relations, Mittelberg (2008; 2010a) and Mittelberg and Waugh (2009) draw upon Jakobson’s theory of two distinct poles of conceptual operations inherent not only in artistic genres (film, literature) but also as figures of thought in general: “similarity (all types of equivalence) and contiguity (temporal and spatial neighbourhood)” (Waugh & Monville-Burston 1990). Jakobson’s distinction accords to Peirce’ fundamental semiotic distinction of similarity (icon) and contiguity (metonymy), which are, according to Dirven, “the two fundamental possibilities of structuring human ‘behaviour’”—in which ‘behaviour’ can be replaced “by ‘conceptualisation’” (Dirven 2003:76).

Of specific interest for this dissertation are the spatial variants of these general semiotic relations: spatial equivalence or similarity as discussed above, and spatial contiguity, henceforth adjacency, both present in Mittelberg’s analyses of metonymy and metaphor in gesture (e.g., Mittelberg & Waugh 2009). Since interaction with objects entails the hand to be outside but in contact with the object, external metonymy and its driving operation of adjacency are vital to the conceptualisation of gestures that imitate these interactions.

According to Jakobson, the semiotic relation of internal metonymy, “the interior aspect of contiguity” (Jakobson & Pomorska 1983:134) describes how a characteristic part or attribute of a concept or an object can stand for the entire concept or object in a \textit{pars-pro-toto} relationship.

\textsuperscript{14} Since gestural representation spans objects that are more or less similar to the hand (or another articulator), this mode covers a larger space in the portrayed continuum.
Internal metonymy in gesture is the relation between articulator and reference object, in which the articulator stands for a part of the referent (Mittelberg 2008; Mittelberg & Waugh 2009), that is, part of its form. For example, wiggling index and middle finger usually do not simply represent walking legs, but stand for an entire walking person (Figure 2.3). In this case, the form of the articulator maps onto the referent only partially, while still referring to the whole object in a *pars-pro-toto* relation, or a relation of internal metonymy. If the wiggling fingers are to represent someone walking, the rest of the walking body is not immediately represented by the fingers. Iconicity works together with metonymic relations in order to make our bodily action stand for a large variety of objects, or abstract shapes such as lines, triangles, and more complex forms (Mittelberg 2010a).

In gestures relying on external metonymy, the hands are external but adjacent to the referent as for example in holding an imaginary object between the hands (Figure 2.7). Here, the hand is external to the imagined object. The hand does not in itself stand for the object, nor part of it. A different kind of metonymy is at work: the imagined object is external and adjacent to the hand, it can be held between the hands or the fingers of one hand, or inside of a cupped hand or on top of an open hand. Any of these examples partially rely on external metonymy and cannot be described exhaustively without this trope (Mittelberg 2010a). Also in drawing, as for example in outlining the shape of an object, we find an example of external metonymy. The hand, again, does not share features with the reference object but rather outlines salient parts of the object’s contour; always remaining at the border of the object, being adjacent to it. In pointing, we find yet another spatial relation that involves external metonymy. Though pointing gestures are able to point at an exact location in space, they can just as well indicate an object whose proportions by far exceed the exact location pointed at. Pointing anywhere on a large table can indicate the whole table, since the table is adjacent to the location pointed at. According to Peirce, contiguity (immediate vicinity, also called adjacency) “is inherent to the index, deictic categories, and, as Jakobson (1956, 1966) also emphasized, metonymy” (Mittelberg & Waugh 2009:334). Following these insights on gesture and semiotic notions implying certain spatial relation, gesture analyses in Chapter 8 will treat examples of adjacency in gestures that point to, draw, and hold (imaginary) objects.

Jakobson further offers another distinction within the notion of contiguity. He calls metaphor and metonymy “artistic transformations” (Jakobson & Pomorska 1983:129) that occur not only in language, but also in other arts, for example in painting and film. Since painting and film are highly visual forms of art, his understanding of contiguity is also applicable to gestures. Jakobson praises Charlie Chaplin’s work because it “broke with the cliches of
proximity in space and mechanical succession in time in order to reassemble the contiguous
associations in original and unexpected ways” (ibid.). Figure 2.5 shows such an “unexpected”
example of a body being spatially contiguous to an object (Chaplin’s body curved around a
large gearwheel). It is exactly this kind of contiguity that I want to use in a Gesture Form
Analysis, distinguishing it from the temporal contiguity (“succession in time”) or the “semantic
proximity” (e.g., of brother and sister; Jakobson 1971:354) by using the term ‘adjacency’: two
forms in immediate spatial vicinity that are shaped in a way that allows for continuous contact.

Figure 2.5: Chaplin’s body curved adjacent to a gearwheel (Modern Times, Chaplin 1936; screenshot JH)

Gestures that include imagining an object adjacent to the hand reveal the object’s shape
because the object can be imagined in tactile contact to the hand or another articulator. For
gesture, it is important to note that metaphor and metonymy do not merge into a single
cognitive process. Following Jakobson, they are dissimilar, even contrasting, notions that
interact in language and elsewhere. Metaphor and metonymy also interact in gesture
conceptualisation as successive steps of interpretation, as suggested by Mittelberg and Waugh.

[I]n order to infer the imaginary objects or traces that gesturing hands seem to hold or draw in the
air, a metonymic mapping between hand (source) and imaginary object (target) is a prerequisite
for the metaphorical mapping between that very object (source) and the abstract idea (target) it
represents.

Mittelberg & Waugh 2009

For the question of how gesture form is created in different gesture types, internal and external
metonymy supply a fundamental distinctions. A model that tries to specify spatial relations
that draw upon these two kinds of metonymy is proposed in Chapter 7. Further, the
understanding of similarity and adjacency interacting in a sequence of interpretational steps will be implemented.

2.5. Mandel’s “sketching” as a systematic analysis of gesture types

Mandel provides insights into how a model of gesture form could look like in the seminal article “Iconic Devices in American Sign Language” (1977). In parts of sign language called classifier constructions, iconic mappings are very prominent, and these constructions have structural similarities to the gesture types discussed so far. Mandel’s theoretical contribution will be applied to the study of gesture form. The proposed model of gesture form can be seen as a framework accommodating his matrix. Mandel’s work underscores the applicability of a model governed by specific spatial operations that can be used to study bodily action. Müller’s “acting” and “representing” can be seen to create form in the same way as Mandel’s “mime” and “substitutive depiction”, respectively. Relying on both scholars, these two gesture types will here be specified as ‘representing in three dimensions’, when the articulator stands for itself as in miming bodily action, or when the articulator stands for something else (fingers representing walking legs).

More important for this dissertation is Mandel’s proposal as to how drawing and moulding systematically create form. For him, these two gesture types are part of a structurally related group, which he calls sketching:

A pointed implement (pencil or fingertip) leaves a line as its trace, either straight or curved depending on the motion. A linear implement (piece of chalk held flat, or a straight or bent finger) leaves a surface, whose shape depends on the shape of the implement (straight or curved) and on the motion. A linear implement can also move parallel to itself, e.g., a straight finger moving like an arrow to depict a straight line. A surface implement (trowel or ice-cream scoop, a whole hand held in various flat or curved surface shapes) can leave a trace shaped like itself, when moved parallel to itself—flat hand moved flat, curved hand moved as if along a curved surface—or it can leave a solid trace perpendicular to itself, as if one indicated the height of a stack of papers by putting his hand palm down on the table and then raising it to the height of the stack.

Under the term sketching, Mandel introduces a matrix that includes the hand both as a drawing tool as well as a moulding tool, adding motion as a factor that generally elaborates the created form. Drawing is a gesture type that uses the “pointed implement” plus motion to create a line. We also recognise moulding in this quote as a “surface implement” that is moved...
parallel to its own extension, thus creating a larger surface. This description does not only analyse the different forms that are active in such a gesture; it shows that, with the creating and the created form, there are multiple forms in one single gesture. It also specifies how motion impacts the created form. The temporal dimension of motion can be translated into an additional spatial dimension. A form of any dimensions is of one additional dimension when interpreted as leaving a trace. This allows for the systematisation of acknowledged gesture types like drawing and moulding and their integration into a matrix of different kinds of form and motion. For example, when a surface implement is moved non-parallel to its own extension, the surface implement creates a solid trace. An example would be someone sweeping laterally with a flat vertical hand to refer to the entire volume of space in front of him (as in Figure 8.58). Mittelberg (2010:378) identifies the core transformation of trace leaving as a spatial metaphor: “in the gesture modality form may become motion and motion may become form (FORM IS MOTION, see also Lakoff and Turner 1989).”

Mandel implicitly takes gesture analysis a step deeper, into the structure of gesture form. On this lower level of analysis, there is not simply an inventory of gesture types but rather a matrix of forms that assemble them in an organized, modular fashion. Pointed, linear, and surface tools can create forms depending on the direction of motion, just as a piece of chalk, lying flat on a planar surface, leaves a different trace when moved parallel or non-parallel to its own extension. This point is illustrated further in Section 3.3 and a systematic account of how these forms interact will be developed in Part II.

2.6. Conventionality, similarity, metonymy, and metaphor in orchestration

For some sign linguists, Mandel’s ideas may appear to challenge the arbitrariness of signed languages, especially since Sign Linguistics fought for the recognition of signed languages as fully-blown languages for some decades (see Wilcox 2004 for a review). Mandel allays worries of this kind by clarifying that iconicity and conventionality are not mutually exclusive.15

A gesture can communicate either by being or making a picture (in a broad sense) or by having a meaning agreed upon by convention, or by a combination of the two. (Most gestures of ASL are combinations.)

Mandel 1977:60

15 For a more detailed account of conventionalisation in gesture and sign language, see also Brentari et al. (2012).
Not only in gesture studies (e.g., Poggi 2008) but also in the more general perspective of
semiotics, the interaction of similarity and conventionality in a single semiotic sign is a given
(Waugh & Monville-Burston 1990; Peirce 1998/1894). Peirce, according to Ransdell, states that

... when we identify some sign as being iconic, for example, this only means that the iconicity of
that sign happens to be of peculiar importance to us for some reason or other implicit in the
situation and purpose of that analysis, but there is no implication to the effect that it is therefore
non-symbolic or non-indexical.

Ransdell 1986:57

Gesture and sign language merely exploit the two semiotic relations in a different way. One of
Mandel’s new insights is to specify a minimum of iconicity and conventionality for gesture and
sign. Regardless the dominance of one of the two semiotic relations, any bodily action has to
contain, in sum, enough iconic resemblance and conventional meaning, to possess a
communicative function. Figure 2.6 illustrates this relation in form of a graph reformating
Mandel’s illustration from 1977.

Figure 2.6: Iconicity and conventionality in signs (redrawn based on Mandel 1977:60; axis direction changed).

Iconicity and conventionality are axes on a shared coordinate system. “100% lexical,
morphemic” signs that bear “no iconic relationship to [the] referent”, as BLACK (in American
Sign Language: moving the index-finger pointing hand past the head at the level of the eye
brows) are part of the inventory of a signer just as “completely ad hoc” body movements, in
which the “form of [the] gesture is identical to [the] form of [the] referent” as in miming ‘I looked at my hand’. The category of the sign BASEBALL combines these two semiotic relations. Although it mimes “hefting the bat at the shoulder”, it is also a fully conventionalised lexical sign (Mandel 1977:59-60). They could be called ‘iconventional’ since they combine iconicity and conventionality. The coordinate system in Figure 2.6 suggests that iconicity and conventionality are the aspects of an action that, both on their own and in combination, afford communication. “A gesture with neither picture nor convention cannot communicate, and is therefore unavailable for use in a language” (Mandel 1977:60). The division of the area between the two axes illustrates this point. The coloured upper right triangle contains all the different actions that are used in sign or gesture whereas movements assigned to the blank lower left triangle cannot be used. Body movements that are neither iconic nor conventionalised do not communicate. Thus, examples of low iconicity and low conventionality do not exist (Figure 2.6). Equally for gesture: bodily action that is neither iconic nor carries conventionalised meaning cannot be a gesture, it remains bodily action only. Wilcox agrees with Mandel in “that arbitrariness and iconicity can be simultaneously present” in gesture and sign language (Wilcox 2004:140). Reaction-time experiments give evidence for iconicity accelerating the processing of well-known signs presented to signers (Grote & Linz 2003).

The coexistence of iconic and conventional aspects in signing might also be reflected in Mandel’s use of “gesture” and “sign”. He sees the mime ‘I looked at my hand’ as a gesture although it is a valid part of signing. Sign language use includes gestures. This is not to equalise gesture and sign, but to account for the similarities among the two. In sign language, we see how gestural action can be a (more iconic and less conventional) part of a full-blown language. Sketching is part of signing as much as it is a part of co-speech gesture. Non-signers, conversely, not only perform idiosyncratic gestures, but also use conventionalised emblems and other gestures.

For the following analyses, also the relation between iconicity and metaphor has to be clarified. McNeill’s (1992) gesture classification, including iconics and metaphorics, has led some to conclude that a gesture is either one or the other—despite McNeill himself proposing that “most gestures are multifaceted—iconicity is combined with deixis, deixis is combined with metaphoricity, and so forth” (McNeill 2005:38). Again, we see that iconicity and metaphor work hand in hand (e.g., Bergmann 2012; Cienki 2008:7; see also e.g., Taub 2004 for sign language): to understand the gesture reproduced in Figure 2.7, which accompanies the spoken utterance that “it was a Silvester and Tweety cartoon” (McNeill 1992: 14) as part of a cartoon retelling, one has to comprehend that the hands pretend to hold or enclose something. They
are moved to a specific position with a specific hand shape, in which they look similar to actually enclosing something between the hands. “The gesture creates and displays this object and places it into an act of offering” (ibid.). If observers fail to understand this, they cannot understand the metaphor IMAGINARY OBJECT is THE NEW CARTOON TOPIC in the discourse. The metaphor builds on the similarity of this gesture to the act of holding something. Without this similarity relationship, the gesture could not be understood. In Taub’s words, “metaphor lets iconic signs have abstract meanings” (2004:3).

Hitherto, this section reviewed findings that showed conventionality and similarity interacting in gesture, and similarity interacting with metaphor. Since a gesture similar to that in Figure 2.7 is frequently used and thus most likely at least partly conventionalised, we can close the circle of interacting notions by adding that conventionality does not exclude metaphor.

Yet, another cognitive-semiotic relation of the above-reviewed literature is missing to account even for a common gesture as that in Figure 2.7. Following McNeill’s interpretation, a “bounded, supportable, spatially localizable, physical object” (McNeill 1992:14-5) is being presented here. One looks in vain for similarities of the hand with this object. Instead, the hands metonymically stand for an object. The cited “physical object” is exactly where the hands are not, it is external and adjacent to the hands, in this case partially enclosed by both articulators (Mittelberg & Waugh 2009). The hands bind the object and spatially localize it,
exhibiting one of its spatial features: the object’s width is defined by the distance between the inner surfaces of the hands.

Summing up, Figure 2.7 is an example of a conventional gesture that is iconic in imitating holding a large object, metonymic in enclosing the imaginary object on two sides, and metaphoric in that the imaginary object stands for the discourse topic (Tweety and Silvester cartoon). All these semiotic relations interact, in fact they build up on each order (“metonymy first, metaphor second”, Mittelberg & Waugh 2009.). While the reviewed literature investigates different aspects of the complex interaction of these conceptual processes in language use, questions regarding the specifics of this interaction (their order, their occurrence across gesture types, etc.) are still open.

In the application to gesture, at least the relations of similarity, internal and external metonymy seem to describe spatial relations that could be captured by geometric relations. It seems that gesture studies can profit from a specific adaptation of these semiotic relations to the motor-visual modality of gestures that take place in three-dimensional space. Put more generally by Hockett (cover-quote of this dissertation), Grote and Linz specify for sign language what also fuels the gesture research in this dissertation:

spoken words are articulated one after another in a linear sequence and stay always in a temporal relation to each other, whereas signs are not only expressed in a sequential order (as in spoken languages) but simultaneously in space as well. Space is directly used to linguistically express properties of a referent like shape, location, motion, manner, direction, features or qualities by movement of the hands through syntactic space. ... Sign languages offer a very interesting field for the study of iconicity because the visually based linguistic system shows a much greater disposition to iconic signs than the auditory system ... Due to the predominance of visual over auditory perception in the interaction with external objects, there are many more possibilities to depict visual similarities than there are to produce acoustic ones in the process of sign-creation.

Grote & Linz 2003:23

Grote and Linz’ view is compatible with research that investigates gesture imitating practices that act on objects, which was the theoretical starting point of this dissertation (Section 2.2).
Focussing on the spatial constraints of semiotic relations—spatialising semiotic relations\textsuperscript{16}— and inquiring how geometric processes are orchestrated in creating gesture form, independent of conventionalisation and metaphoric mapping, is the goal of Gesture Form Analysis.

\textsuperscript{16} What does this ‘spatialising’ specifically entail? The variety of notions that this dissertation builds on are not necessarily spatial in nature, but those that do have clear spatial implications will be boiled down exactly to those spatial implications. Let us consider external metonymy as an example. There is one kind of external metonymy whose spatial realisation is rather straightforward. Given the hands are external but adjacent (Mittelberg & Waugh 2009) to the imaginary object as in Figure 2.7, the spatial constraints of external metonymy are clear: the object extends from one hand to the other (see also Sowa 2006). One can therefore measure the width of the object by applying the concept of external metonymy to the position of the two hands (and if Figure 2.7 showed someone telling what big fish he caught, this would even be a useful computation). The spatial relation between hands and object is clear; the semiotic relation can be broken down to spatial (geometrical) formulas, such as

\[ width_{\text{imaginary object}} = \text{distance (hand\_right, hand\_left)}. \]

But external metonymy is a semiotic relation also applicable in ways, for which it is difficult or impossible to state a specific spatial relation. It also applies to relations, in which “the name of a referent is replaced by the name of an attribute, or of an entity related in some semantic way (e.g., cause and effect; instrument; source)” (Wales 2001:252, capitals removed) as in the classic example: “The ham sandwich is waiting for its check.” (Lakoff & Johnson 1980:35). The spatial vicinity of the ham sandwich and the customer in the restaurant is not in the foreground, but rather that the customer has ordered (or eaten) the ham sandwich. Here, we have a causal or instrumental relation, which could not be captured satisfactorily by a spatial formula as in the prior case. This dissertation focusses on those semiotic relations that have spatial implications and attempts to organise these implications in one framework (see Chapter 7).
3. Forms in gesture literature

The series of Chapters 3 to 6 will explore the areas of gesture studies, human perception and (Sign) Linguistics with regard to the structure of how spatial information is conceptualised. Across the diverse areas relevant to the perception of the phenomenon of gesture, the conveyed pieces of spatial information are described in a similar way: the descriptions contain forms of specific dimensions. Some gesture types are described by forms of zero and one dimension (by points and lines); other types are described by forms of two dimensions and three dimensions (by surfaces and volumes). Specific form dimensions seem to be a characteristic of gesture types and the cognitive processes of their perception. While gesture analyses typically imply different types to be related to specific forms, only few accounts specify form dimensions as a criterion that characterises certain groups of gestures (Müller 1998; Mandel 1977; Sowa 2006). A framework that accounts for the systematic occurrence of specific form dimensions in different gesture types does not yet exist, as far as I know.

Beginning with gesture studies itself, Chapter 3 will investigate descriptions of a variety of gesture types and extract the use of terms that specify forms and specific form dimensions. The analysis of forms will show that different gesture types share some forms while diverging in others. These commonalities and differences in gesture form will be taken as the core structure of a gesture typology as developed in Part II of this dissertation.

3.1. Observing bodily action

As already suggested by the selected literature, the investigation of gesture form is intended to be viewed from an embodied perspective.

Embodiment in the field of cognitive science refers to understanding the role of an agent’s own body in its everyday, situated cognition. For example, how do our bodies influence the ways we think and speak?

Gibbs 2005:1

‘... and express ourselves,’ might be added more generally. Gesture is a part of communication that, described systematically, is a good subject matter to investigate specifically “the role of an agent’s own body” as part of language use. Gestures will be investigated as a phenomenon that is action as well as utterance (Kendon 2004).
While traditional disciplines such as Psychology, Linguistics or semiotics offer distinctions and categorisations that we can apply to gesture, this work aims to contribute to gesture-specific theory. I will focus on gesture as a special type of (manual) action in space. Other action strategies such as manipulating objects or exploring shapes can have similarities to the types of action we perform when communicating. Everyday life is defined by the manipulating power of the hands in two ways. Firstly, with the hands, we touch, carry, push, pet, crush or use any kind of tool. Each of these actions is a well-trained, complex motor program. Secondly, we observe others perform all these activities and understanding the goal of these actions is also of utmost importance. The correct interpretation of these activities relies on our ability to identify patterns in complex actions when we see them. Both the knowledge of how to perform an action and the ability to identify action patterns in others are likely to be reflected in gesture conceptualisation. The basis for understanding a gesture is the capacity to visually perceive the body of the interlocutor, held still or in motion, and recognise patterns: patterns which allow inferring imaginary objects or other meaningful forms. The relation of both parts of gesture perception, seeing body movements and inferring specific forms from them, is the perspective taken in this dissertation. Gesture Form Analysis describes processes of gesture form interpretation that are necessarily embodied, since gesture form is a derivative of how, in everyday action, the body moves and how this movement is interpreted.

When performing ordinary actions as well as when gesturing, we usually also see, feel and hear what we are doing at the same time. In this sense, the actor is at the same time an observer. Perceiving one’s own action is often an important part of the action itself. Carrying out an action without seeing it can feel very different. Also, precisely performing ‘handling an object’ in the absence of the object is difficult. Otherwise, pantomime would not be an art but something anyone can do without practice. For example, visual and tactile observation is a constant source for adaptation of the action. This dissertation deals with forms relevant for both the production and the perception of gesture. Still, the study focusses slightly on the observer. Note that while only the gesturer can be called the producer of the gesture, both the gesturer an onlooker can be called observers because they perceive the gesture and act on the basis of this perception. The perspective of the gesturer, the perspective of the interlocutor, and the perspective of any other by-stander is different, but they are all observers of the gesture. For example, both the person pointing in a specific way (Kendon 2004:110-4) as well as the person perceiving this pointing gesture can examine where the gestural articulators are pointing to—the gesturer with the purpose of verifying correct
pointing and the observer with the purpose of identifying the referent. The next chapter reviews the literature on pointing in order to identify the forms that are necessary to fulfil both of these purposes, referent identification and referent verification.

### 3.2. Pointing contains different forms

Some detailed analyses of pointing describe how exactly articulatory action has to be conceptualised in order to yield the information associated with pointing, be to single out a referent or to indicate a direction. Sometimes more, sometimes less implicitly, scholars name specific spatial dimensions that will be treated as a distinctive feature in the model of Gesture Form Analysis developed here.

The first gesture type whose analyses will be reviewed with regard to form dimensions is pointing. As one of the clearest examples of action that is understood as communicative, gestures of pointing, or deictic gestures, have been recognized as a separate class by almost all of the students of gesture we have reviewed and it has always been understood that such gestures can play a fundamental role in establishing how an utterance is to be understood. There are very few studies, however, which have examined the way in which pointing is done.

Pointing, or the group of indexical gestures, is widely treated as a proper gesture type. However, the relation of pointing to other gesture types differs (Kendon 1988a, 2004; Mandel 1977; McNeill 1992; see Enfield 2003 for an overview). In other words, it is not agreed upon which other gesture types are similar to pointing. Motivated by the agreement that pointing is a powerful and ubiquitous part of gesturing and the disagreement with regard to its relation to other types, I would like to closely examine “the way in which pointing is done” (quote above).

Pointing gestures have a very clear form, without which indicating something outside of the body, sometimes even very distant, would not be possible. Its gesture form sets it apart from other gestures. It is also gesture form, as will be argued here, that shows that pointing

---

Vector calculation isn’t as trivial as it first appears.

*H. H. Clark 2003:252*

Kendon 2004:199

---

*Kendon 2004:199*
shares structural elements with other gesture types and where these structural overlaps lie. Kendon and Versante’s assessment of how a pointing gesture succeeds in singling out an object alludes to structural aspects relevant to gesture form:

pointing gestures are regarded as indicating an object or a location that is discovered by projecting a straight line from the furthest point of the body part that has been extended outward.

Kendon & Versante, 2003:112

In this description, we encounter terms that indicate gesture form. Apart from the verbal root “point” in “pointing”, Kendon and Versante mention a “straight line” and a “point” in their definition of a pointing gesture. Lines and points could just be one arbitrary way among many to talk about how we point. In this dissertation they are seen as characterising and structuring the gesture type pointing; they are forms without which pointing could not be successful. Favouring this perspective, these terms reoccur when people are asked to explain pointing in detail.

3.2.1. Line or vector

This section argues that a careful analysis of pointing contains specific kinds of forms: those of a point, and a line or vector. Although few scholars who use these terms focus on geometrical implications, it seems inevitable to refer to concepts that overlap widely with those of point and line when talking about pointing. Still, even the scholars that specify the geometry of pointing gestures usually do not illustrate these dimensional forms for purposes of visualisation—likely because their existence as part of identifying the referent seems self-explanatory. Exceptions are found in computational approaches, for example Pfeiffer (2010) visualising a vector (dark-grey line) being part of pointing (Figure 3.1).

17 To avoid confusion, I will not use the noun ‘point’ to stand for ‘pointing’, but always to stand for a ‘point in space’.
Kita opens his edited volume *Pointing: Where Language, Culture and Cognition Meet* by stating that a pointing gesture at least must contain a “vector”—regardless of the “further types of signs [it] creates” (Kita 2003a:1-2). Pointing is said to entail a “vector” in research from different disciplines such as Linguistics (*i.a.*, Bühler 2000/1934:79,125 (“Pfeil” = arrow); Fricke 2007; Streeck 2009), Sign Linguistics (*i.a.*, Liddell 2003), Psychology (*i.a.*, Butterworth 2003), Ethnology (*i.a.*, Haviland 2003; Wilkins 2003; Enfield 2009), and Computer Science (*i.a.*, Kranstedt *et al.* 2006; Pfeiffer 2010). Even though these works span a wide range of approaches and the use and theoretical implementation of the specific term ‘vector’ varies, there is at least a minimal agreement among these areas: even diverging definitions of a vector agree on it being a linear form that has a specific orientation in space indicating a direction.

In pointing, people “appear to be extrapolating a vector through space […] based somehow on the angular orientation of the gesture or the movement of the pointing arm” (Butterworth 2003:24). “[A] body part projects a vector toward a particular direction” (Kita 2003a:5; see also Enfield, Kita & de Ruiter 2007). Clark (1996; 2003) further states that in “principle, speakers can exploit any body part with which they can create a vector” by treating the articulator’s “major axis as a vector” (2003:251). This vector neither seems to be bound to a specific articulator, nor to the static shape of the articulator: one can even “compute the vector of the head’s back and forth motion and follow that vector from the body” (*ibid.*).

In analysing Arrernte speakers, Wilkins (2003:184) identifies a group of “hand signs [that] are deictic and communicate vector in an analog fashion.” The term analogue reflects the ability to continuously rotate the articulator, thereby continuously rotating the vector

---

For a critical discussion of the „vector hypothesis“, see Butterworth 1997 2003:23 *et seqq.*
associated with it. For a form analysis, a vector is almost the same as a line. As a line (or, more precisely a line segment), a vector has a certain length or magnitude. Besides, just as a line, a vector is defined in its orientation (or rotation) in space. A further characteristic of a vector is that it contains, in addition to mere rotation, the extra information of direction: "the direction of what is being pointed at is seen as away from the pointing hand" (Butterworth 2003:10). Other literature speaks of a line that is being "marked" in space (Fricke 2007:105) or that the forefinger can "constitute" a line (Calbris 1990:128).

A vector, or a line plus direction, is constitutional part of the gesture type pointing. A model of pointing has to contain the notion of a vector or a line. Part of the gesture form of pointing is a line.

Gesture types that also contain lines may be candidates for structural relatives of pointing. However, Gesture Form Analysis does not regard a gesture type simply as an unordered sum of forms. As will be shown in the Section 3.3, lines can be part of a gesture type in different ways, or, more precisely, the different forms that make up each gesture type are ordered in a particular way and build up on each other (Part II, Section 7.2).

3.2.2. Point (intersection at a certain location)

As mentioned initially, the line is not the only form examined frequently in analyses of pointing. At least for a certain variant of pointing, ‘pointing at’, the vector through space is used “to intersect with a potential target” (Butterworth 2003:24). Clark (2003:253-4) claims that in pointing, one does not indicate an entity directly but rather a “conceptually conspicuous site”: what it being indicated “is not a person or object, but a location”. This could be interpreted as that even the gesture that points at an entity, first points at a point in space. Fricke (2007:105) pinpoints this idea: pointing at something is accomplished by a “point in space” (Raumpunkt) that, in turn, “localises” an entity.

Is the contiguity relation identified in the immediate contact of the articulator and an object (Mittelberg & Waugh 2009; Talmy 2012) equivalent to the contact of the point and an object, as conceptual constituent of ‘pointing at’? Talmy’s general notion of a “connector construct” that includes “contiguity” (2012:6) in order to gesturally single out a distant referent indicates such a structure.

3.2.3. Relation of forms in pointing

Is it possible to systematically distinguish different kinds of pointing on the basis of the geometric forms that are necessary in “an object or a location [to be] discovered” (Kendon&
Versante 2003, quote beginning of this chapter)? The literature does not exactly approach the problem of gesture categorisation from this perspective, but Kita names three kinds of things pointing can draw attention to, which I interpret as implying a structural hierarchy. Pointing can direct the attention of someone to “a certain direction, location, or object” (Kita 2003a:3). I apply these different targets of pointing to a form analysis in proposing that drawing attention to a direction, a location and an object, each requires the gesture’s conceptualisation to include a different set of forms. What Kita does not state explicitly is that pointing to the latter two entities, “location” and “object” includes the first, “a certain direction”. We cannot point at a location without pointing in a direction; and we cannot point at an object without pointing in a direction and pointing at a location. For each of these three spatial entities, an observer might need to interpret the given articulator form (e.g., the index-finger pointing hand) one step further.

This is consonant with Fricke’s (1) “direction deixis,” (“Richtungsdeixis”) and (2) “point-in-space deixis,” (“Raumpunktdeixis”, 2007:105-117, translations JH). She classifies gestures that (1) indicate a direction as different from gestures that (2) indicate a point in space or a “self-contained entity” associated with that point in space. It is not made explicit how the “point” is localised in space but I assume that the point is a product of the intersection of two geometrical forms: first, the line that extends from the articulator and second the surface of the entity pointed at. The intersection of forms, like the line and the surface that produce a point, will be investigated systematically across gesture types in the search for another general geometric operation of gesture form (Part II, Section 7.2). For now, summarising the overlapping implications in Kita (2003a) and Fricke (2007), Figure 3.2 shows a basic form hierarchy of two kinds of pointing, directing and pointing at location.

<table>
<thead>
<tr>
<th>&quot;line&quot; extending from articulator</th>
<th>plus: &quot;point&quot; at intersection</th>
</tr>
</thead>
<tbody>
<tr>
<td>&gt; directing</td>
<td>&gt; pointing at a location</td>
</tr>
</tbody>
</table>

Figure 3.2: Kinds of simple pointing gestures.

The same forms also seem to be part of pointing in signed languages. As argued in detail by Johnston (2013), “pointing signs found in signed languages are not fundamentally different from the pointing actions found in the composite utterances of spoken languages” (ibid., p.
109), including the differentiation of pointing “to help identify a location” and “pointing to help identify an entity” (ibid., p. 110). Differentiating object identification from the kinds of pointing in Figure 3.2, as implied by Kita (2003a) and Johnston (2013), will be addressed at a later stage (Section 8.4).

3.3. Drawing and pointing on common ground

Gestures called “drawing”, used similarly by various scholars (Goldin-Meadow, Kim & Singer 1999:721; Mandel 1977; Mittelberg 2008; 2010a; Müller 1998; 2013; Poggi 2008:50; Williams 2008:69;), are widely seen as a gesture type independent of pointing. It will be shown that presupposing a categorical difference between drawing and pointing does not consider the constitutive similarities between the two gesture types (as also suggested in the work of Goodwin 2003a; Haviland 2003; Mittelberg & Waugh 2009). One apparent indication of structural similarity is the articulatory equivalence between what could be called a pointing-hand and a drawing-hand: for example an extended index finger, other fingers curled in. More generally, any articulator that protrudes visibly from the rest (or even an elongated object) can be used to both point and also to draw. Both types, pointing and drawing, therefore, can overlap in articulator form. They make use of similar hand shapes. What distinguishes the two is that drawing creates a line in the air, more precisely a path or trajectory. As will be shown, in terms of gesture form, drawing a line in the air heavily relies on the forms that constitute the gesture type of pointing.

Are the gesture types pointing and drawing related by the forms that are part of their conceptualisation? If so, implications of their common structure, with regard to gesture form, should be found in the literature. Indeed, Kita summarises parts of Haviland’s work on pointing (2003) by saying that “a pointing gesture can create an iconic representation by tracing a shape or movement trajectory” (Kita 2003a:2). I would like to adhere to this idea and investigate their relation more closely. Are the forms that were found to be part of ‘pointing at a location’ (Figure 3.2) also part of a drawing gesture? The following description by Haviland suggests such a structural overlap.

“Moreover, in addition to the ... punctual extension of an outstretched limb, other sorts of formatives, including motion, accompany apparent pointing gestures.”

Haviland 2003:161
The controlled motion within the stroke phase is treated as a potential supplement to a pointing gesture that turns it into a drawing gesture. I propose pointing and drawing to overlap in their structure of gesture form, similar to the structural overlap of ‘directing’ and ‘pointing at a location’ (Figure 3.2). These gesture types are a representative portion of a gesture typology as proposed in the end of Chapter 8. Understanding both a pointing and a drawing gesture includes imagining certain forms: a line that extends away from the body and a point where this line intersects with a surface. This is a structural overlap with regard to gesture form. Drawing includes not only these forms but also an additional form that sets it apart from pointing and that highlights its additional, more richly iconic, function: the line that is being drawn by moving the hand (and the form that it conceptually creates). Goodwin supports this claim from another perspective. Also criticising general distinctions of deictic and iconic gestures, he opts more generally for different components:

In most typologies of gesture (see McNeill, 1992, p. 76, for a summary), iconic gestures and deictic (pointing) gestures are treated as separate kinds of gesture. This does not seem to be correct. Pointing gestures can trace the shape of what is being pointed at, and thus superimpose an iconic display on a deictic point within the performance of a single gesture. Instead of using this distinction to separate gestures into distinct classes, it seems more fruitful to focus analysis on an indexical component or an iconic component of a gesture, either or both of which may contribute to the organization of a particular gesture.

Goodwin 2003a:229-230; see also 2003b

Goodwin’s and Haviland’s take on the interaction of iconic and indexical “components” comes close to McNeill’s (2005:42) work referring to “dimensions” of iconicity and indexicality. Investigations into metonymy (Mittelberg 2008; 2010b) and metaphor (Müller 2007) elaborates on this interactional view of semiotic or spatial relations. For the structure of gesture form, regarding iconicity and indexicality as “components” that occur in a distinct succession (Mittelberg & Waugh 2009) is particularly relevant.

In consonance with Goodwin, Haviland and Mittelberg, Mandel takes the description of the interaction of shape and motion to another level—though preceding these scholars’ work by a quarter century. He identifies the pointing-drawing relation as a structural feature among

---

19 It should be noted that McNeill is using the term “dimensions” in the sense of scalar properties. In this dissertation, we limit the use of dimensions to refer to spatial dimensions.
these and other gesture types and thus supports the need to search for a multi-layered system of forms in how gestural action is conceptualised. Mandel’s contribution to the issue is not a single relation but a consistent matrix, as laid out in the longer text quote above (see Section 2.5; Mandel 1977:67). He introduces motion as a factor that interacts with other forms in gesture. Hand shape and motion (depending on the direction) are both factors influencing gesture form, their conceptual product. He calls the group of gesture types in which motion is conceptualised this way “sketching”. Sketching includes but is not limited to drawing. Not only points, but also lines and surfaces can leave traces. For lines and surfaces leaving traces, direction is crucial (Figure 3.3). Motion parallel to the shape’s own axes leaves traces similar to its own form. A line leaves the trace of a longer line (Figure 3.3a) and a surface leaves a trace of a larger surface (c). Contrastingly, perpendicular motion lets a line (e.g., piece of chalk held flat) leave a surface as its trace (b) and lets a surface leave a volume as its trace (d). Mandel’s use of “perpendicular” should be extended to non-parallel, by which then all possible movements are covered in binary distinctions. This would then include diagonals (b) and cases where the motion is neither parallel nor perpendicular, but any angle in between.

As the concept of drawing suggests, some gestures using this technique can be imagined to draw on a surface, just like physically drawing with a pen on a sheet of paper or with the finger in the sand. Müller offers a clear example of this in a participant describing a “course of a love relation as ‘a kind of up and down’” (Müller 2008b:234). In a discourse segment characterised by focal attention on the gesturing hand and foregrounded metaphorical expression, the participant draws a curvy line from left to right which gradually decreases in magnitude and descends. Meanwhile, she says that the relationship “began like this and then flattened like this on out” (“startete so und flachte dann so weiter ab”, ibid.; see Müller (2008b:233-8) for illustration and analysis). The drawn line stands for the quality or intensity of a relationship, as

Figure 3.3(a-d): Trace leaving of different shapes in different directions. Lines (a,b) and surfaces (c,d) in parallel (a,c) and non-parallel (b,d) motion.
in a graph, so that this example is empirical evidence for drawing gestures being open to metaphorical interpretation.

For Müller, drawing gestures result in a “two dimensional line drawing” (Müller 2014:3689) and the gesture on the ‘love-relationship’ shows this being an intuitive characterisation. Nonetheless, gestural drawing does not have to be imagined on a two-dimensional surface. Drawing can also employ depth, thus exceeding two-dimensional space. For example when “a person who has been asked what a winding staircase looks like describes with his finger a rising spiral” (Arnheim 1974:48), this helix-formed trajectory clearly does not adhere to a two-dimensional surface to draw upon. Further, the curvy run of a cable on a construction site can be depicted in gesture space without adhering to any surface. Drawing seems to be variable and not necessarily bound to a surface to draw on. Further, drawing can also be less complex, not requiring a surface. Simply drawing a straight line in the air exhibits nothing but a one-dimensional trace. Imagining this line on a surface is possible but by no means a necessary characteristic of this gesture type. Whether more or less complex, the line as a product of a drawing gesture is not two-dimensional, but follows Mandel’s more general trace-leaving rule of a point leaving a line as its trace.

Also Calbris implicitly uses motion to substitute for an additional spatial dimension. She says that a circle can be represented by the index finger and the thumb held as if enclosing a circle or the index-finger pointing hand performing a circular motion: the “circular shape may be represented either statically or dynamically” (1990:50). The static circular shape of index and thumb in itself (Figure 3.4a) or a point at the tip of the finger in motion leaving a circular trace (Figure 3.4b) can equally represent the shape of a ‘circle’.

For Katz (1925), this principle is also found in active tactile perception. The movement itself, which is performed to explore the shape of an object by attaining different tactile sensations
while sliding over and adapting to an object’s surface, is cancelled out from the conceptual result of the tactile perception.

As we will see, the sense of touch characteristically reacts to successive stimulation with structures to which nothing of motion in any form appears to belong. It is as if the cinematic form of the stimulus is converted to the static properties of an object.

Katz 1989:79

The same is true for visual perception in that although the image of the room on the retina shifts when someone rotates the head while looking around, it is still perceived to be static. It is just an “optical motion” (Arnheim 1974:379) caused by active visual exploration that is not actually perceived as motion. Following Katz and Arnheim, perceiving certain kinds of motion as a means of exploring a shape, and not as motion as such, seems to be a general cognitive and perceptual process that is more general than gesture perception. It is as well an integral part of concepts in spoken and written language, as Talmy highlights.

Motion as an additional dimension of form can be seen as a variant of Talmy’s “fictive motion”. “This fence goes from the plateau to the valley” contains fictive motion because the fence is factively not moving, but language uses terms characterising a motion event (‘go’ ‘from’, ‘to’) to describe this static configuration (Talmy 2000:99). Fictive motion in a language is characterised not only by something static being described as something dynamic, but also by the fact that the modality of speech or written language does not carry this dynamicity in itself, but merely refers to a dynamic concept by specific terms and constructions. Gestures describe static configuration with motion, too, only that ‘fictive’ motion can be expressed in the modality itself. It may still be called fictive since the mapping’s target (the fence) does not move, but this fictive motion is expressed by factively moving the hands. Another aspect of fictivity then lies in a part of the hand leaving a trace (that is to be mapped onto the shape of the fence). Mandel’s term of “atemporal motion” (Mandel 1977:66–7) expresses this notion of trace leaving as a movement (extending through time and space), from which the aspect of time is extracted resulting in a motionless shape: the trace. Thus, in the gesture research conducted in this dissertation, I expect to encounter many instances of this kind of ‘fictive’ motion.

Summarising some structural aspects of the relation between pointing and drawing gathered in the literature, both types may be integrated into one hierarchical structure as
illustrated in Figure 3.5. In this figure, every gesture type contains all the forms of all the types to its left. Directing and ‘pointing at a location’ must contain the concept of a line. Moreover, pointing at a location contains a point where the line intersects with the surface of an (imaginary) object. Drawing, then, is another elaboration, in which the point of intersection is moved and leaves a trace, as depicted in Figure 0.1 (Introduction). Whereas a pointing gesture could be called iconic with regard to location, since the point indicates the position of a referent, drawing also reveals the shape contour of the referent. By means of example, Figure 3.6 shows an archaeologist (“Ann”) drawing an imaginary curved line on the ground some centimetres below her finger. The author (Goodwin 2003b) himself added an arrow to mark the imaginary trace created by this gesture.

<table>
<thead>
<tr>
<th>&quot;line&quot; extending from articulator</th>
<th>plus: &quot;point&quot; at intersection</th>
<th>plus: moving, the point &quot;leaves a trace&quot;</th>
</tr>
</thead>
<tbody>
<tr>
<td>&gt; directing</td>
<td>&gt; pointing at a location</td>
<td>&gt; drawing</td>
</tr>
</tbody>
</table>

Figure 3.5: Drawing gestures as ‘pointing gestures that leave a trace’.

Figure 3.6: Drawing. See Figure 8.24 for a Gesture Form Analysis of this example. From Goodwin (The Body in Action (2003b:21), Palgrave Macmillan; excerpt). Reproduced with permission of Charles Goodwin and Palgrave Macmillan. This material may not be copied or reproduced without permission from Palgrave Macmillan.

3.4. Holding and moulding: surface-contact gestures

In the very rich data of human movement analysis, such as gesture research, one could expect certain simple geometric forms to occur in the descriptions of all kinds of gestures. But this is not the case. Whereas lines and points are used frequently to describe pointing and drawing, they are seldom used for describing other types of gestures, for example in the gestures Müller calls moulding. In these gestures, the hand traces over the surface of an imagined object,
thereby indicating the shape of an object. In this technique, according to Müller, the hands form a three-dimensional gestalt (Müller 1998:118; see also Müller 2010:177), see Figure 2.2c. Besides the movement characterising the shape of the referent, the surfaces of the hands can indicate the proportions of the imaginary object (see also Mittelberg 2010). The imagined reference object has to be imagined in contact, more specifically in surface contact with the hands. It will be investigated how adjacency or "touch", which for Jakobson is one of the key features of external metonymy (Jakobson & Pomorska 1982:114; Mittelberg & Waugh 2009), plays out in the specific spatial relation of surface contact in gesture.

In order to understand and differentiate what is meant by the technical term of ‘surface contact’, a definition of different types of contact is in order. Surface contact is to mean the contact of two objects that are contiguous in their shape, for example a cupped hand partially enclosing a baseball. This contrasts to the kind of contact between a pencil touching a tabletop with its sharpened tip. All kinds of pointed shapes, but also a round shape, as a billiard ball on a table, are understood to touch a flat rigid surface at a point, rather than touching a surface area (at least on a human-scale level of detail). Roughly, also the tip of the finger is understood to touch a flat surface at a point. They are in ‘point contact’. A knife, with the cutting edge of the blade oriented downwards, touches a flat, horizontal surface in a line, as, roughly, also does the edge of the flat hand against a table top. They are said to be in ‘line contact’.

Surface contact occurs often when at least one of the two touching objects is non-rigid, such as a sponge, a mass of gum, or, more importantly, the tactile surfaces of the hands when they touch an object. This definition of surface contact overlaps with Gibbs’ use of “surface contact” (2005:213). It deviates from the broader use of “surface contact” by Croft (2009:$32), Streeck (2009:149), and Levinson & Wilkins (2006:520), which encompasses the contact of objects of any shape being placed on a surface—regardless of whether the two objects touch at a point, along a line, or all over a surface area.

Apart from moulding, there is another prominent group of gestures that implies surface contact: holding gestures. One example of holding is the palm-up-open-hand gesture, which is often used in the (metaphorical) function of offering something to the interlocutor: the “Palm Up Open Hand presents an abstract discursive object as an inspectable one—an object which is concrete, manipulable, and visible [...]' (Müller 2004:233). For Gesture Form Analysis, it is not trivial that the imaginary object is not the hand itself but rather to be imagined on the hand, adjacent to it. It is because of this adjacency that Mittelberg sees the “palm-up open-
hand gestures with a flat palm or cupped hand evoking a kind of surface or a receptacle where items can be placed \((i.e., \text{imagined})\) and presented to the audience" (Mittelberg 2010a:360). The example from McNeill (above, Figure 2.7) is another gesture whose conceptualisation involves imagining surface contact.

For the gestures that hold or mould an object, at least two forms seem to be relevant in the analysis of their gesture form: the surface of the hands and the solid shape of the imaginary object. This partly deviates from Müller and Sowa who attribute only one number of dimensions to gestures that fall into this category. According to these authors, these gestures are three-dimensional (Müller 1998) or either one-, two, or three-dimensional (Sowa 2006).\(^{20}\) I agree with the cited authors in that holding and moulding gestures can depict a three-dimensional imaginary object (Müller), that the (two-dimensional) surface of the articulator being in contact with the surface of the object is a crucial form interpreting such a gesture (Mittelberg), and that holding gestures as in Figure 2.7 can indicate the one-dimensional extent of an imaginary object (Sowa), but I would like to propose a framework that accounts for the coexistence of all of these forms as part of the conceptualisation of holding and moulding gestures (Chapter 7). A closer investigation is required of how a surface (in contrast to a line, being a constituent form of pointing and drawing) affords a range of functions such as containment (Mittelberg 2010a) or maybe even more generally: division of volumes. Figure 3.7 summarises the surface-contact gestures of holding and moulding, characterised by the articulator being adjacent to the imaginary object.

<table>
<thead>
<tr>
<th>tactile 'surfaces' of the hands enclosing an imaginary object</th>
<th>plus: hands moving to show the shape of the imaginary object</th>
</tr>
</thead>
<tbody>
<tr>
<td>&gt; holding</td>
<td>&gt; moulding</td>
</tr>
</tbody>
</table>

Figure 3.7: Surface-contact gestures.

\(^{20}\) Sowa would classify McNeill’s example (Figure 2.7) of holding an imaginary object as one-dimensional since it is “indicating a one-dimensional extent” \((\text{ibid., p. 95, 253})\).
3.5. Representing in three dimensions

The articulator can also itself represent the referent in its three-dimensional shape. This gesture type is henceforth referred to as ‘Representing in three dimensions’. Liddell, in his analysis of ASL, has generalized this concept to apply to any articulator, including the entire body. In his conception, the articulator becomes a “surrogate” that is conceptualised within the given scene.21

If the signer’s body becomes someone or something else through blending, I describe the resulting blended space as a *surrogate space* and describe the elements within that space as *surrogates*.

Liddell 2003:152

3.5.1. Mapping all or just selected form aspects

In Representing in three dimensions, some or all aspects of the complex shape of the hand can be relevant for the form (of an object) that it represents. Mandel (1977) gives the example of signing ‘I look at my hand’ by actually looking at one’s hand. Here, the hand stands for the referent in its entire, complex, and of course three-dimensional, form. Looking at one’s hand does not have to be a gesture—it could also just be an action—but it can be (part of) an utterance, and thus a gesture. It can be part of a scene, in which the signer or gesturer depicts herself to be ‘looking at her own hand’, or depicts someone else who is not the narrator ‘looking at her own hand’. In both cases, the displayed articulator form can be mapped onto the referent hand shape in its entire form (*e.g.*, configuration, position and movement). Note that the form parameters here are acceptable to describe gesture form—in contrast to all the preceding examples—because gesture form, in this special case, is limited to articulator form.

In distinction to these very rich mappings, there are many examples in the sign language classifier system and also in gesture, in which the articulator represents the referent in its three-dimensional features, but via a selective mapping; for example in signing that someone walks by wiggling index and middle finger alternately, called BIPED-WALK (Figure 2.3 and Liddell 2003, *e.g.*, 265 and 305). The fingers—and only the fingers, and not the rest of the hand or arm—stand for walking legs: in being longish, having a certain thickness, being able to bend

---

21 For an extensive analysis of surrogates in ASL, see Liddell (2003). Surrogates, for Liddell, are created within specific types of blended mental spaces (see Section 6.2 of this dissertation and also Williams (2008a) for more on blending).
backwards at what is conceptualised as waist and knee. But there are also many features that are not mapped: humans have only one knee per leg unlike the two joints of a finger; a leg does not have a big nail at its lower front, but rather a foot with five toes, and so on. The example is here taken from sign language, and there are many others, such as the index finger (pointing upwards) standing for a (literally) upright person (*ibid.*, 270) or a certain hand configuration standing for a vehicle including its three-dimensional orientation (*ibid.*, 276).

Of course, variants of BIPED-WALK occur in gesture just as well (for gesture examples, see, e.g., Calbris 2011:15; Duncan 2002:199). Representations like the BIPED-WALK are still three-dimensional but they cannot be as rich as the ‘I look at my hand’, exactly because of the partial matching of features. Another gesture example of this kind is using the hand instantiating a telephone. Calbris (1990:105) examines that “a French person mimes, in a single gesture, the grip of the caller (digits 2 3 4 folded on the palm), [and] the shape of the receiver (digits 1 5 spread).”

3.5.2. Different articulators represent

As in the ASL sign BIPED-WALK, in gesture, often only a (three-dimensional) portion of the gesturing body is used for the mapping. That different parts of the body are mapped onto different things is again scrutinised for sign language by Dudis as “body partitioning” (2004) and is paralleled by the analysis of “multiarticulatory gestures” spanning the simultaneous analysis of multiple body parts (Hassemer 2009). An example of multiarticulation would be the hands illustrating the shape of an (abstract) referent at the same time as the eyebrows marking important segments of the utterance and the gaze pointing to certain locations that are associated with already-established referents (Hassemer 2009:57-74). Communicative functions and gesture types are not limited to specific parts of the body. Rather, function and type are articulator-independent. Several parts of the body can perform gestures: any finger, the hands, arms, shoulders but also the head (Kendon 1990), the lips (Enfield 2009), the eyes (Hassemer 2009), among others (e.g., see list of articulators in Clark 1996:166). Gesture examples of this dissertation (e.g., Figure 8.45 and Figure 8.6) are to underscore that, in principle, any gesture type is articulator-independent (though the hands are very flexible and thus convenient for several types).

3.5.3. Representing and cognitive iconicity

In general, among gestures in which the articulators stand (partially or entirely) for something in its volumetric shape, we can make one simple distinction regarding the form of the
articulator in relation to the form of the entity that the articulator is representing: the articulator can stand for itself (thus being very similar or equal to the referent) or it can stand for something else (being less similar). ‘Representing in three dimensions’ spans the entire similarity continuum (introduced in Section 2.3, see Figure 2.4).

When the articulator represents itself, or another hand, its shape and movement overlaps largely with what it stands for. But when it represents something else, in Lakoff and Johnson’s terms, there are two “different kinds of things” (1980:5), there is only a perceived similarity of the articulator with the object. This similarity is not about physical, objective features only. The hand with pinkie and thumb extended in a ‘telephone gesture’ is not similar to a receiver of a telephone. The hand is still a hand, consisting of flesh, blood and bones that have nothing in common with a plastic receiver filled with cables, membranes and air. Hand and receiver can also not be used for similar things: Making a call with the hand is as impossible as tying a knot with a telephone receiver. The two things are different. Only the proportions of an extended hand and a telephone receiver are vaguely similar; only their form is minimally iconic. The convention of representing a telephone with this hand shape is one cue for this three-dimensional mapping. Further, our perception of the hand performing a gesture that is part of a verbal-gestural utterance makes it possible for the hand to represent a receiver. The specific context of use as part of an utterance, in which a telephone is a possible referent, as well as the location and orientation of where the hand is held and the displayed facial expression, makes a comparison possible. Only the complete orchestration of the bodily action in addition to the discourse context makes it possible to speak of similarity. The hand and the receiver are iconic within the framework of “cognitive iconicity” (Wilcox 2004). Cognitive iconicity holds that two entities such as hand and telephone receiver are similar only in their representation in “conceptual space” (Wilcox, Wilcox & Jarque 2003), in other words how they are conceptualised, or, in Langacker’s (2008) terminology, “construed”.

---

[C]ognitive iconicity is defined not as a relation between the form of a sign and what it refers to in the real world, but as a relation between two conceptual spaces ...

Iconicity is not a relation between the objective properties of a situation and the objective properties of articulators. Rather, the iconic relation is between construals of real-world scenes and construals of form.

Wilcox 2004:122-3
Wilcox’ perspective can be traced back to Peirce’ study of meaning involving not only the sign and the object, but also an interpretant “(whatever a sign creates insofar as it stands for an object)” in the observer’s mind (Kockelman 2013:165). In contrast to Saussure, Peirce’ theory focusses on meaning being “motivated and context-bound”, foregrounding “inferential relations between signs and interpretants” (ibid.). It is the similarity of two interpretants, the conceptualisation of the gesture with the conceptualisation of the object, which drives the perception of an iconic gesture.

3.6. Different kinds of form in a single gesture

Except for the one group of gestures treated in the preceding section, the gesturing hand often stands for a vector as in pointing and drawing, or for a surface when surface contact to the object is to be imagined. Given the fact that we often talk about physical, three-dimensional objects and our articulators are also three-dimensional, it could surprise that we do not use our articulators more often to represent with their full height, width and depth.

There are two reasons for the ‘conceptual detour’ of using additional low-dimensional forms (line, surface) as intermediate steps to refer to physical objects or abstract shapes. One is that the shape of the hand is rarely similar to the shape of the object. This argument alone is not convincing given how minimal a similarity serves to tie conventionalised gestural action to a referent such as in BIPED-WALKING, PERSON-UPRIGHT, or the telephone receiver discussed above.

A stronger argument is that acting as the object itself is not part of our normal repertoire of every-day action. The parts of our body can be directed towards objects (reaching out for something) or in contact with objects (holding something), but never spatially overlap with objects. In playing basketball, the hand never stands for the basketball. It touches, pushes, holds or throws it. The hand is in surface contact with the ball, it never is the ball, and it never is in the ball. Real-world interactions accustom us to manipulate imagined objects as being external, but adjacent to the body. This is not only a metonymy in the sense of a cognitive or linguistic phenomenon, but a spatial design feature of human action with objects in general. These spatial structures of everyday interaction will be taken as a starting point for the principles of gesture form. This is not to neglect the differences there are between the surface contact of a hand holding an imaginary object and the surface contact of a hand actually holding an object. The tactile sensation is clearly different in both cases and especially a lax gestural performance will be non-stable in its surface contact with the object. Of relevance for the approach at hand is the equality in the spatial structure of action on a real object and
gesturing with an imaginary object. This spatial structure is defined by the material of the hand and of the object being in surface contact.

The perspective taken includes the relation of structural elements of gestural action in particular and action in general. The adjacency of objects to our body (surface contact) and the directedness of our gaze and also other body parts, toward desired entities (concept of a vector; see also Talmy 2000: Chapter 6) might, both in action and gesture, be a source rather than a consequence of conceptual and linguistic structure. Within an embodied perspective that emphasises “the role of an agent’s own body in its everyday, situated cognition” (Gibbs 2005:1), specifically the “world at hand” (Streeck 2009:58) being guided by “embodied schemas” (Mittelberg 2010a), Gesture Form Analysis takes structures of every-day action as candidates for conceptual principles of how humans convey spatial information in gesture. This is not to equate holding a ball with gesturing to hold a ball. Clearly, the performance and also the perception thereof are different in the two cases. Yet both in the factual case of holding an object and in the gesture imitating holding, a three-dimensional object (the hand) contacts another three-dimensional object (ball/imaginary ball) and the spatial extension of the contact is a surface (surface contact). With regard to the structure of the involved forms, action and gesture are claimed to show a significant overlap.

Reviewing gesture analyses of the gesture types pointing, drawing, holding and moulding as well as three-dimensional representations has shown two structural aspects that will shape the further investigation: (1) gesture scholars’ descriptions include specific forms for some of these categories, such as ‘line’ and ‘point’ in pointing and drawing. (2) Even ‘simple’ gestures, such as pointing, seem to consist of more than one form; at least a vector or line in addition to a point at the intersection of the line with the surface of the entity pointed at. These two findings motivate the theoretical efforts in Chapter 7. It will be examined whether there is a fixed set of differentiable kinds of form, and whether the strategies discussed in the preceding sections can be re-analysed in terms of these kinds of form. A scrutiny of gesture types with regard to gesture form requires a structure of analysis that is based on general principles and specific operations, both abstracted from spatial relations of everyday interaction with objects. On this level, a model of gesture form conceptualisation and a gesture typology is conceivable.

To be able to break gesture types down into clearly distinguished forms for a lower level of analysis, distinctions between these different kinds of forms have to be grounded in differences of form and function. Potential functional implications, or affordances, of the basic forms encountered to be part of gesture types are those of (1) most ‘efficient’ (meaning least complex
form that fulfils a specific function) and precise indication of position (for the zero-dimensional point), (2) most efficient and precise indication of direction (for the one-dimensional line), (3) most efficient and precise division of volumes of space (for the two-dimensional surface) and (4) most efficient and precise description of an object’s shape and movement (for the three-dimensional volume).

How the articulators in a gesture, themselves complex three-dimensional entities, stand for different lower-dimensional forms will be addressed explicitly in Part II. Since we cannot see most of the spoken languages’ articulators (vocal apparatus), gesture studies and Sign Linguistics covers new ground within the spatial domain of body-movement interpretation by dealing with the perception of the visible articulators, for example the hands.

(A)nyzing iconicity requires that we examine not just our conceptualization of objects and events in the world, but also of articulations—hands and movements.

(Wilcox 2004:124)

Chapters 4 to 6 try to locate the analysis of form within the bigger picture of perception and interaction on the one hand, and low-level structure of (sign) language on the other hand. This is for the purpose of connecting the concepts of gesture form with the concepts of perceiving and understanding bodily action in general and using spatial relations for describing spatial communication in particular.
4. Forms in perception

This chapter defines a particular understanding of abstraction, and introduces a phenomenological view on perception. The continuum of perception and conception already touched upon in the Introduction will be embedded in the Gestalt Psychological perspective put forth by Arnheim (1974; 2004/1969) and Gibson (1977).

It will be argued, that reception of signals in itself is widely two-dimensional, at least for the gesture-relevant senses of vision (via the retina surface inside the eye) and touch (via the tactiley sensitive surface of our body, the skin). In the further processing, especially of visual information, surface information remains of predominant importance as shown for example in Marr’s 2½D sketch (2010/1982), a processing stage that produces specific kinds of surface information only.

This chapter encompasses a basic review of perceptual structures that are part of making sense of the world around us, including people’s gestures. We usually have to see in order to interpret a gesture. An account of gesture perception and conceptualisation must be consonant with general perception, on which it relies. This chapter is an attempt to attend to this interdisciplinary demand.

4.1. Abstraction

Arnheim’s definition of abstraction in relation to gesture lays the ground for claims on gesture perception in general and gesture form in particular: gestures communicate spatial information that is much more reduced, or more abstract, than the complex motoric action by which it is conveyed.

Actually, the portrayal of an object by gesture rarely involves more than some one isolated quality or dimension, the large or small size of the thing, the hourglass shape of a woman, the sharpness or indefiniteness of an outline. By the very nature of the medium of gesture, the representation is highly abstract. What matters for our purpose is how common, how satisfying and useful this sort of visual description is nevertheless. In fact, it is useful not in spite of its sparseness but because of it. Often a gesture is so striking because it singles out one feature relevant to the discourse. It leaves to the context the task of identifying the referent: the bigness portrayed by the gesture can be that of a huge Christmas parcel received from a wealthy uncle or that of a fish caught last Sunday. The gesture limits itself intelligently to emphasizing what matters.

(Arnheim 2004/1969:117)
In this quote, Arnheim illuminates the way gestures communicate; in a perspective also assumed in this dissertation. He emphasises the power of abstraction in gestural communication. Gestures convey information about the referent in a highly abstract fashion. Further, incorporating Wilcox’ above-cited stress on “conceptualization [...] of articulations—hand and movement” (2004:124)—the form of the gesture is also abstracted away from the rich spatial information of the hand or another articulator (e.g., their configuration and orientation). A certain hand shape (two flat hands opposing each other) might trigger the observer to recognise a gesture to stand for the width of an object only. What is conveyed by the gesture is so abstract that, depending on the context, this width can be mapped onto very different objects including Arnheim’s examples of a parcel or a fish.

To be able to assume Arnheim’s viewpoint on gesture, his definition of abstraction has to be clarified further, in order to see which among the many possible interpretations of this term’s wide and heterogeneous use is meant.

The crudest misuse of the two terms, then, is that of saying ‘concrete’ when ‘perceivable’ is intended, and ‘abstract’ to describe what is not accessible to the senses.
It is equally misleading to call concrete that which is physical and abstract that which is mental...
Any phenomenon experienced by the mind can acquire abstractness if it is seen as a distillate of something more complex.

Arnheim 2004/1969:155-6

Arnheim deals here with a fundamental problem of abstraction. Among a set of diverse objects, humans are able to identify an abstract concept that discriminates one set of objects from the rest. But how do we generate this abstract concept? It is necessary to generalise this concept over all the objects of a set to be able to tell whether the abstract concept suits all of them; and none of the objects outside of this set. “This means that an abstract concept, supposed to be the fruit of generalization, turns out to be its necessary prerequisite” (Arnheim 2004/1969:159). Do we just pick any one feature and assess whether an intended object happens to share it? This procedure could not account for the quick performance in humans in recognising a shape never seen before to belong, for example, to the concept of spheres.22

22 Even different animals recognise simple shapes, for example when they associate them with a treat (Sutherland 1960; 1969; Baldwin 1981).
To shed light into the sequential order of abstraction and generalisation, the choice of the abstracted concept has to be motivated. To specify this motivation and in his enterprise of discussing border regions of perception and thinking, Arnheim draws on Bergson, using his “utilitarian origin of sense perception”. Perception is seen “as a means of discovering the presence of what is needed for survival and for being alerted to danger” (Arnheim 2004/1969:160). The flywheel of abstraction would then be the Darwinist urge to group our surroundings for example into ‘dangerous’ and ‘not dangerous’ to accomplish survival. Put more generally and in Koffka's words:

To primitive man each thing says what it is and what he ought to do with it: a fruit says "Eat me"; water says, "Drink me"; thunder says, "Fear me," and woman says, "Love me."

Koffka 1935:7

In a general Gestalt Psychologist’s view, we perceive objects with regard to how they might be useful to us. This motivation is not limited to ensuring survival, but also any other kind of benefit or any one step in accomplishing an every-day goal. If we widen the criterion of survival to identifying how something is ‘helpful’ or ‘productive’, Arnheim's understanding of perception and abstraction is in line with Gibson’s theory of affordances. Gibson quotes another passage of Koffka’s “Principles of Gestalt Psychology” (1935) highlighting the applicability of affordances to every-day action.

The postbox "invites" the mailing of a letter, the handle "wants to be grasped," and things "tell us what to do with them [p. 353].” Hence they had what Koffka called "demand character."

Gibson 1977:77

Affordances, for Gibson, relate to “geometrical shapes” and Section 4.2 will exemplify what the specific geometric shapes of two-dimensional “surfaces and their layouts afford” (1977:72). The theory of affordances links form and function on a very basic level. Perceiving an object and making sense of it is one and the same (complex) cognitive process. Other approaches of Cognitive Linguistics and gesture studies subscribe to this view (i.a., Talmy 2000; Streeck 2009).
Talmy notes that there “is no principled way to draw a line between perceiving and conceiving” (Talmy 2010:160). Cognitive operations relevant to the approach at hand also belong to “processes that don’t properly belong to just perception or conception, but in fact extend across a continuum comprised of the two of them” (ibid.). Talmy calls the synthesis of both processes “ception”:

it seems advisable to establish a theoretical framework that does not imply discrete categories and clearly located boundaries, and that recognizes a cognitive domain encompassing traditional notions of both perception and conception. Such a framework would then further allow for the positing of certain cognitive parameters that extend continuously through the larger domain (as described later). To this end, we adopt the notion of *ception* ...

Talmy 2000:139; highlight in original

In this dissertation, I will use perceiving and perception as well as conceiving and conceptualisation as being located on Talmy’s continuum of *ception*—with the tendency of perception to focus on early perception and conceptualisation to focus on ensuing cognitive operations. The notion of image schemas as coined by Mark Johnson (1987; for applications in gesture, see, *e.g.*, Cienki 2005 and Mittelberg 2010a) might be an example of minimal structures that are located in the centre of the continuum between perception and conceptualisation.

4.2. Surfaces in perception

The remainder of this chapter deals with two kinds of perception that are part of making sense of bodily action: we perceive of gestures visually, while gestures imitate actions that include tactile sensation. Indeed, there is a range of work describing how gestures are related to actions involving objects (*e.g.*, LeBaron & Streeck 2000; Müller 1998; Streeck 2009; Wundt 1904). The central sensory system involved in manipulating or exploring different kinds of objects is touch. In general, sight and touch are the dominant senses for the perception of space. “Skin and retina are, however, the organs in which the space-element plays the most active part” (James 1890:135). Surface information is characteristic of these two gesture-relevant kinds of perception, visual as well as tactile. The stimulus signal is processed by receptors that are distributed in a certain pattern across the body: many of these receptors are arranged

---

23 Streeck (2009:9) uses the same term for a somewhat similar but not equivalent notion, the details of which are not part of the argument at hand.
inside of thin surfaces—the visual receptors on the retina and the tactile receptors on the skin. While these receptors are also distributed at different depths within these thin surfaces, the depth information itself is not utilised. In visual perception, the front sensors do not specialise in receiving stimuli that are farther away, nor, in tactile perception, do the rear cutaneous sensors specialise in receiving stimuli penetrating the skin.\textsuperscript{24} Contrastingly, different positions on the surface of receptors are distinguished. In vision, the location of the receptors that receive visual information is important, and, in tactile perception, at which point on the surface of our skin we feel pain (see, e.g., Van de Graff 2001).

The processed signals feed into our concepts of the objects in our surroundings. In an extreme position, the perception of objects can be broken down to "an arrangement of simple geometric components, such as blocks, cylinders, wedges and cones" (Biederman 1987:115). I will follow this reductionist approach in attempting to identify specific form categories in visual and tactile perception that play an important role in perceiving gestures. Further, aspects of established theories on visual perception will be exploited. Also in a non-reductionist perspective on visual and tactile perception, one can maintain that the form of an object is largely perceived by its surface and not immediately by its solid three-dimensional form\textsuperscript{25} (see below; e.g., Marr 2010/1982; Ellis & Young 1996), as will be introduced with the help of the subsequent example.

Looking at a metal sphere, for instance, we cannot tell whether it is solid or hollow. Visual perception, including stereovision and active exploration (by moving the object around and looking at it from different perspectives), does not allow us to penetrate the surface of an object to look inside. It only informs us in detail about the location, colour and other characteristics of, and only of, its surface. Also by touching a metal sphere we will obtain information about its surface only. We infer the extent of the object through the curvature of the surface that encloses the object, but still do not know whether it is solid or hollow. We can infer that the sphere is solid by different interpretations of surface information. When we hold the sphere in one hand, we can infer its weight from the sensation of the vertical force that the surface of the object exerts on our hand. On the basis of this downward push being strong, we could conclude that the object is solid. However, it could just as well be a predominantly...

\textsuperscript{24} Rather the kind of signal determines receptor position, since, e.g., receptors for deep pressure (organs of Ruffini, lamellated corpuscles) are located in a lower layer than receptors for light pressure (corpuscle of touch, bulbs of Krause) (Van de Graff 2001:490-1).

\textsuperscript{25} As in earlier examples, the dimensions of the object are meant. Obviously, the two-dimensional surfaces are located in three-dimensional space.
hollow sphere containing a small lead core. A surface that quickly assumes the temperature of our hand may favour interpreting the sphere as hollow because we know from experience that a hollow object like an empty tin is rapidly warmed by the contact of our hand. Still, in the case of the lead-core sphere, also this conclusion would be erroneous.

Visual and tactile perception of surface information is not the only aspect of a structure of everyday perception and interaction that can be applied to how we perceive gestures (Streeck 2009). Understanding body movements as gestures should involve looking for patterns of goal-directed action with objects as well as strategies, for example, of tactile exploration of objects. “Our fixed motor patterns correspond to affordances of objects which these reveal during our daily, repetitive tasks” (Streeck 2009:56). For the question of gesture form, Streeck’s approach raises the question of whether gestures display not only affordances of objects that the gesturer imitates to be interacting with, but whether the forms reoccurring in descriptions of certain gestures (points, lines, surfaces), have general affordances that characterise the types they are part of.

For example, spanning over iconic and metaphoric mappings, the large group of gestures that interact with imaginary objects exhibits a specific pattern: the hand’s (inner) surface is in contact with the imaginary object, which means that it can exhibit parts of the imaginary object’s shape, more precisely only surface characteristics, for example a certain curvature. This central role of surfaces is paralleled by their central role in the general perception of our surroundings. It is the set of geometrical forms that can be summarised under the term surfaces, which define our concept of a solid object’s shape. No matter what the shape of a solid object may be, its surface, by definition, encloses it. Conversely, surfaces constitute the shape of a solid object in its entirety. Put, more generally, surfaces divide space: to one side of the surface is only the object and nothing else, and to the other side of the surface is everything else and nothing of the object. Surfaces determine an object’s shape and the shape of an object is one of the characteristics determining what we can do with the object; in other words, what the object affords. Put in Hutchins’ terms, within the concept of ‘material anchors’, “the association of conceptual structure with material structure can stabilize conceptual representations” and surfaces are part of “the complexity of the material structure” (Hutchins 2005:1555).

In introducing his theory of affordances, Gibson even identifies surfaces with the material they enclose: “the ‘shapes’ of surfaces [...]”, by which is meant the solid geometrical shapes”. This is to be understood in the framework of what he calls ecological physics, in which he
conflates properties of substance and surface. He claims that these “physical-geometrical features [...] do not belong to separate realms of discourse; they are one and the same. Geometry [...] is connected with life!” (Gibson 1977:72). This affordance-based perspective will be assumed in the proposed Gesture Form Analysis. Also, surfaces will be treated in their special function of delimiting material and thus dividing space. Our body and especially our hands afford changing the world on the basis of surface contact with our surroundings.

Since we do not smell, taste and seldom hear gestures, visual and tactile perception are per se more closely related to performing and understanding gesture. Visual perception is also the key sense for the perception of space, be it the shapes objects or the spatial relations of objects, including measuring the ‘empty’ space between them. Haptic perception, apart from also exploring peripersonal space, is active, too, when we perform the physical actions that certain gestures imitate: holding up an object (Kendon 2004; Müller 2004), sliding the hand over an object’s surface to explore it (Müller 1998; 2010; 2014) or brushing aside undesired objects with the back of our hand (see Müller & Cienki 2009:304-6). These are examples of contiguity relations and external metonymy in everyday action (see also Mittelberg 2013). The empirical study included here (Chapter 9) will show that the participants haptically explore, hold and visually observe physical objects before describing them. This chapter’s focus on surfaces is due to their importance in perception and actions. Moreover, the gesture typology will prominently feature a category of surface contact gestures that a wide range of the analysed data will fall into (Section 8.5).

4.2.1. Perception as active exploration

The approach developed in this thesis adheres to a phenomenological and Gestalt Psychological perspective on perception, which is often adopted in Cognitive Linguistics. This is expressed for example by Arnheim’s contention

that the cognitive operations called thinking are not the privilege of mental processes above and beyond perception but the essential ingredients of perception itself. I am referring to such operations as active exploration, selection, grasping of essentials, simplification, abstraction, analysis and synthesis, completion ... These operations are not the prerogative of any one mental function; they are the manner in which the minds of both man and animal treat cognitive material at any level. There is no basic difference in this respect between what happens when a person looks at the world directly and when he sits with his eyes closed and "thinks". ... I must extend the meaning of the terms 'cognitive' and 'cognition' to include perception.

Arnheim 2004/1969:13
“Active exploration”—in Arnheim’s terminology clearly belonging to the domain of both perception and cognition—is an example of haptic action that is imitated in moulding and drawing gestures. Though naming several mental operations that highlight the interplay between active exploration and thinking, the structure of this interplay is not specified by Arnheim. Are there general cognitive operations of abstraction, for example in reducing the complexity of the form of an articulator movement down to meaningful and communicative constituent forms?

If the terms that Arnheim uses to describe cognitive operations really are “the manner in which the minds of both man and animal treat cognitive material at any level”, then the way in which we conceptualise gesture form should be describable with this very terminology: in making sense of the complex shapes around us, including those of an articulator while gesturing, do we “select the essential”, “simpler” forms as for example the approximate curvature of a hand’s surface? If so, could these form features that we observe in the way how people “grasp”, “actively explore”, or more generally, “analyse” objects in gesture be “completed” or “synthesised” to stand for the surface curvature of a referent object (all citations from the larger text quote above)? These questions on how solid articulators like the hand are systematically abstracted to their tactile surfaces (see e.g., Mittelberg 2010a), will be formalised in Chapter 7 and applied in Gesture Form Analyses of Chapter 8.

4.2.2. Visual perception

This and the following section touch upon the processing of stimulus information in visual and tactile perception in the search of specific forms recurring in the representations of spatial information along different processing stages of vision and touch.

When light reflects off of the surface of an object and enters the eye, it falls onto the surface of the backside of the interior of the eyeball, the retina. The ‘picture’ on our retina is the input for our visual sensation. The raw material of visual perception consists of different degrees of colour and brightness on a flat image, a two-dimensional picture without depth information. The photoreceptor cells are arranged on the retinal surface and capture a two-dimensional projection of the light that enters the eye. Being arranged on a surface is a characteristic of many, but not all of the receptors that contribute to spatial perception. Receptors that relay information about body position and movement are partially distributed throughout joints and muscles (Van de Graaff 2001:488-90). Still, the rod and cone cells in the retina of the eye, as well as the different receptors responsible for touch, pressure and pain on
the skin, belong to the broad class of exteroceptors, distinguished as being “located near the surface of the body, where they respond to stimuli from the external environment” (ibid., 489). The sensory impulses transmitted to the brain from these sources thus do not contain any three-dimensional data.

In Vision (2010/1982), Marr suggests that not only the first 'input picture' is two-dimensional but that also what he calls the following two steps of visual processing, the primal sketch and the 2½D sketch, contain (different kinds of) surface information only. The primal sketch comprises an image, a flat picture, in which intensity changes are recognised, including the “local two-dimensional geometry of an image” (Marr 1980:199). This planar image contains already all the spatial information needed for gestures that convey information that projects onto a frontal plane (see Figure 8.32 and Figure 8.42).

The 2½D sketch, on the other hand, is a map that contains surfaces of objects including their orientation and depth in space. Both the primal sketch and the 2½D sketch rely on two constraints (C1 and C2) of the physical world around us:

(C1) a given point on a physical surface has a unique position in space at any one time; and (C2) matter is cohesive, it is separated into objects, and the surfaces of objects are generally smooth compared with their distance from the viewer.

Marr & Poggio 1979:302

The second constraint (C2) draws attention to an interesting feature of the distribution of mass in our environment. Mass is not scattered in random-sized chunks across three-dimensional space. Objects, materials, and substances tend to be arranged in cohesive shapes that exhibit surfaces that are generally smooth, at least from a human-scale perspective. The visual field is usually not completely filled up with heterogeneous contours and edges but consists of only a limited number of surfaces, whose contours and edges stand out in relation to their comparably smooth surfaces that result in regular, or at least regular gradients, of brightness and colour. Marr includes this very basic characteristic of the visual field as a constraint on visual processing that highlights the importance of surfaces and limits early visual processes to the detection of surfaces and edges (Marr & Hildreth 1980). In the first two stages of visual perception, the reception of light of photoreceptor cells on the retina (input picture) and the first step of visual processing, the primal sketch, both deal exclusively with surfaces. More precisely, they deal with flat images of intensity changes. To understand the
role of surfaces in the next steps of visual perception, a slightly more detailed introduction is necessary.

At first sight, the 2½D sketch, the subsequent step of visual processing, seems to leave the domain of purely two-dimensional surfaces. Clearly, this step goes beyond a single flat image. But what exactly does the ½ dimension in the ‘2½D’ sketch stand for? The 2½D sketch is “something like an ‘orientation and depth map’ of the visible surfaces round a viewer” (Marr & Poggio 1979:306). The ½ dimension stands for one additional kind of information on the visible surfaces identified at this stage. This extra information contains the orientation of surfaces (which Marr & Poggio indicated by small arrows in Figure 4.1) as well as the rough distance between them, in relation to the observer. Different parts of a surface are determined in their orientation in space: that is, where they are facing, as indicated by the small arrows in both parts of the figure. Thus, additional edges, which were not already marked in previous processing (for example, because they emitted similar brightness and colour), can be inferred.

The 2½-D sketch is a dynamic memory with considerable intrinsic computing power. It belongs to early visual processing, and cannot be influenced directly from higher levels, for example via verbal instructions, a priori knowledge or even previous visual experience.

Marr & Poggio 1979:318

Relying now on “a number of different and probably independent processes that interpret disparity, motion, shading, texture, and contour information,” (Marr & Poggio 1979:306) the flat areas of the primal sketch are now curved and bent surfaces and located at different depths in space. Nevertheless, these surfaces clearly do not contain any depth information of the solid
objects they stand for; they are still, slim sheets of space. This means that early visual perception, up to and including the 2½D sketch captures nothing but ever richer information on surfaces. The surface dominance does not stop at this point. Rather, all early visual processes seem to aim at the

recovery of gradually more objective, physical properties about an object’s shape. The main stepping stone toward this goal is describing the geometry of the visible surfaces, since the information encoded in images, for example by stereopsis, shading, texture, contours, or visual motion, is due to a shape’s local surface properties. The objective of many early visual computations is to extract this information.”

Marr 2010/1982:36

The question arises whether all of visual perception is about more than a complex arrangement of surfaces. Marr’s definition of shape suggests that.

I shall reserve the term shape for the geometry of an object’s physical surface...

Marr 2010/1982:296; highlight in original

Marr, as well as Gibson (see above) define an object’s shape as equivalent to the geometry of the object’s surfaces. This equivalence has implications for the different ways in which objects can be represented, for example in gesture: showing the surfaces of an object would thus be equivalent to showing the entire shape of an object.

Following Arnheim in considering perception as a part of thinking, a large part of our thinking about objects is about surfaces only. This means that physical objects, by definition three-dimensional, are processed as different kinds of surfaces for a large part of our perceptual processing. Only in the final stage of the “3-D Model Representation” (Marr 2010/1982:302), surface information gets processed to provide us with a three-dimensional map of our physical environment.

“One must infer the structure of a 3-dimensional world from a 2D projection, a process far beyond simple ‘matching.’”

Ostrovsky 2010:5
Because humans are accustomed by the steps of perception, the inference of three-dimensional objects from two-dimensional surfaces could allow them to infer imaginary objects when surface information is available (Grandhi et al. 2012). For example, the surfaces of the hands when performing actions on imaginary objects can be used to compute the imaginary object’s shape. Part II will elaborate on this idea in terms of a specific gesture form operation (Adjacency (Enclosure), e.g., Sections 7.2.8, 8.5).

One counterargument to calling the different stages of visual perception a kind of ‘surface vision’ could be the depth information that stereovision provides. The two surface pictures of each eye in combination allow computing distances, but these pictures do not immediately capture the distance to the object; they only capture the size and proportions of the object from two different points of view, with which we can infer the amount of empty (three-dimensional) space between object and viewer. Two surface pictures also do not allow capturing the depth of the solid, opaque object: looking at an object with both eyes, we are not able to distinguish a solid object from a well-made, hollow dummy. Summing up, this sketchy account of vision suggests surfaces playing an important role in the perception and conceptualisation of gesture form.

4.2.3. Tactile perception

Tactile perception in general relies on several different kinds of receptors that are spread over the skin and throughout different parts of our muscles (Van de Graff 2001:488-95). This unequal distribution is one of many ways in which visual perception is different from tactile. Another is that we can only perceive things by touch that are close to us, whereas the operating distance of visual perception is not limited; given there are no occlusions. Still, there is one aspect of visual perception discussed above—that of the heavy reliance on surfaces—which it has in common with tactile perception.26

Tactile perception is reduced to the contact of a surface of the body, say, the palm, with the surface of the stimulus, for example the surface of a rigid object. This type of contact will be called surface contact. It follows that the different receptor types, regardless of whether they process temperature, pain, pressure or vibration, are arranged in a way that limits the

26 There are, of course, other aspects of similarity between visual and tactile surface perception, for example that in vision and touch we have areas of much lower perceptory resolution. To give just one example: in peripheral vision, we have difficulties estimating the size of objects (Goodale & Humphrey 2005), as we have difficulties differentiating the location of two nearby stimuli on the shoulder (but not on the tip of the finger; Lederman& Klatzky 2009:1441-2).
attainable spatial information about the object, the localisation of a given stimulus, to what can be obtained from its surface. By touch, for example, we can immediately perceive that a stone is hotter on one part of its surface than on another, but we cannot perceive that the stone is hotter inside than outside.

In the monograph *The world of touch* (1989; orig. *Aufbau der Tastwelt*, 1925), the phenomenological psychologist David Katz stresses surface contact as the only possible form of tactile experience. In the following quote, he also emphasises the core affordance of two-dimensional forms: they enclose space. As generally the case in all material physical entities, every entity is bounded by its surface. Katz’ perspective entails that tactile information is perceived in the form of surfaces, including the spatial exploration of shape and other characteristics of objects.

It doesn’t make any difference, whether we’ve got an adamant material such as glass or a soft woollen fabric stretched out over a solid base. It is always the case that a two-dimensional touch-structure is brought to our consciousness, which encloses space … [E]ven the softest and loosest surface will always bring touch impressions that in their structure are surface palpations.

Katz 1989:43

To analyse surface contact more deeply, we have to differentiate within the broad category of tactile perception. (1) We feel a tactile sensation, when an object is in contact with us, even when we do not move, for example when another person touches us or when we simply put our hand on an object. (2) On the other hand, we sometimes intend to explore an object and thus move our hands over the object in order to perceive an object by touch. Stout (1899) defines the two strategies as follows.

(1) The hand, either open or closed, may touch simultaneously the parts of the object. This may be called *passive touch*, because it does not involve active movement from one part of an object to another… (2) A portion of the hand, such as the fingertips, may explore the parts and contours of the object by gradually moving over them. This may be called *active touch*, because it essentially consists in active movement.

Stout 1899:342-3
The perception strategy (2) is also called haptic perception (Flanagan & Lederman 2001). In haptic perception, following Stout, motion plays a distinctive role in different practices of exploration. I would like to clarify, in terms of surface contact, in what way exactly motion is to be understood to distinguish passive from active touch, as this will serve to distinguish the role motion plays in gesture. It is not the case that passive touch involves no motion and active touch does involve motion. When touching an object in what Stout calls passive touch, we still move our hands to the object to touch it and we will retract the hands afterwards. This is relevant as a gesture imitating this action will most likely include these movements in the preparation and retraction. So the criterion motion per se does not differentiate passive from active touch. What sets these two haptic actions apart is what Stout describes as “[exploring] the parts and contours of the object by gradually moving over them”. In terms of surface contact, this can be stated as active touch being the perceptual strategy in which the surface of the hand moves in relation to the surface of the object. This is what sets it apart from passive touch where the hand’s surface, once touching the object, does not move in relation to the object’s surface. Both of these exploratory movements may be imitated by gesture and the different employment of movement—the presence or the absence of relative motion of the hand’s and the object’s surfaces—is a candidate for a distinction in the structure of gesture conceptualisation.

The condition of static contact (no relative motion of surfaces) does not only characterise the perceptual strategy of passive touch. In grasping any kind of object, we seek to establish enough friction to allow for static contact because that enables us to move an object by moving our hand. The widespread range of examples for this includes many manipulatory actions: from picking up a glass of water to drink from it and poking a sponge, to more complex actions with objects such as unscrewing the cap of a bottle, handling a sword, packing a suitcase or even moulding clay. The manipulatory action strategy of grasping involves establishing static contact between the hand’s and the object’s surface. The object is manipulated because by grasping something we can alter the position of (a part of) the object, thus changing its position (if rigid) or shape (if not rigid). The limitation of passive touch is that it does not allow us to feel an area greater that the hand’s surface.

With active touch, in contrast, we can extend the area of haptic exploration arbitrarily by moving the hand over the object. Thus we can perceive objects that are much larger than our hands and of complex form. Of special interest is the role of motion in this action, as this role is very different from the manipulatory actions just dealt with.
As a matter of fact such movements yield the perception of what may be called geometric configuration, configuration in which the spatial character is not regarded as belonging to any external body, but as the product of our subjective activity. To quote Professor James: "If, with closed eyes, we trace figures in the air with the extended forefinger (the motions may occur from the metacarpal, the wrist, the elbow, or the shoulder-joint indifferently) what we are conscious of in each case, and indeed most acutely conscious of, is the geometric path described by the finger tip. Its angles, its subdivisions, are all as distinctly felt as if seen by the eye; and yet the surface of the fingertip receives no impression at all ... In persons born blind the phenomenon in question is even more perfect than in ourselves" [Principles of Psychology, vol. ii., p. 190].

In these geometric tracings we are making an express experiment, and concentrating attention on the movements of the finger, as such. Under such conditions there is present a very distinct mental image of the path described by the moving finger. It is as if the finger actually left a marked track behind it, and so drew figures in the air.

Stout unintentionally anticipates the relation between active touch in an action and a gesture moving over an imaginary object's surface: when, in a gesture, the traced object is missing, the observer of an imitated tracing movement can interpret this movement as 'trace-leaving', enabling the drawing of a figure in the air (using a finger-tip) or the moulding of objects (using the whole inner surface of the hand and fingers).

In summary, different types of surface contact determine different types of tactile perception and, in the senses most related to gesture perception, those of vision and touch, spatial information results from surface information. In fact, early visual processing, at least up to the 2½D-sketch is, according to Marr (2010/1982), limited to surface information. Surface information in vision and touch is one example of reduced shape representations in cognition and perception of a three-dimensional environment. Shifting from perceptual to conceptual processes, the next chapter will touch upon how grammar constrains the conceptualisation of a referent to two-, one or zero dimensions.
5. Forms in language about space

In this dissertation, it will be attempted to find form primitives for the conceptualisation of gesture. The approach adheres to the idea that perceiving and conceptualising the shape of an object means "fitting the stimulus material with templates of relatively simple shape" (Arnheim 2004/1969:27).

Though building on work on Gestalt Psychology and image schemas, the terms 'gestalt' and 'image schema' will not be used as part of Gesture Form Analysis. This is not in opposition to work on image schemas. On the contrary, the forms that are part of the conceptualisation of gesture, for example, are image schemas. Yet, I have difficulty using this term to productively differentiate the constituents of a gesture, for example into schematic and non-schematic forms—rather, all forms of less than three dimensions are schematic (clearly, also 3D forms can be schematic, as, e.g., the imaginary object in McNeill’s cartoon example in Figure 8.30, being specified only in its width). In describing how the form of a gesture is conceptualised, characterising spatial information with the help of geometric terminology (spatial 'dimensions', 'curved line', 'orthogonal') seems sufficient.

As with 'image schema', the term 'gestalt' is sometimes used in a broad or even vague way. For example, gestures are investigated with regard to "scalar gestalt features: e.g., more rising than falling, more symmetrical than asymmetrical, more checking than flowing. A set of such tendencies constitutes a gestural gestalt for a gesture" (Hirsch 1994:482). The criterion of a gesture being "more rising than falling" is readily operationalisable and does not require the notion of gestalt, which would rather blur an otherwise clear-cut criterion. The other criteria of symmetry and manner of movement would have to be operationalised before being able to introduce them into an analysis that specifically deals with the characteristics of forms. In doing so, the notion of gestalt would give way to a structural description that attempts to capture the interaction of different factors: a model. Both terms sometimes appear to be used as a passe-partout for a complex system whose components and their interdependent relations are not (fully) known. Using the term 'gestalt' does change the need to put forth relations or operations and test their descriptive potential. While the notions of gestalt and image schemas are very general (thus applying to many different concepts) and not necessarily spatial, Gesture
Form Analysis can be understood as a spatial specification of these notions by a model of primitives\(^{27}\) (basic kinds of form) and operations (describing specific form constraints).

For the conceptualisation of gesture in the three-dimensional gesture space, I hope to find specific criteria for gesture categorisation (Part II), and also to be able to formulate a testable hypothesis in a study involving motion-capture data (Part III). In describing how several criteria work together, I would like to propose very basic (in the sense of simple) structural coherences instead of adhering to the rather broad scope of image schemas and gestalts. Candidates for gesture form primitives are spatial dimensions of form as found in grammar (Talmy 2000).

5.1. Talmy’s topological dimensions

As will be laid out in this section, Talmy provides precise distinctions of form conceptualisations in language, while still assuming a Gestaltist perspective on perception and abstraction of form. Speech is usually not immediately ‘spatially iconic’\(^{28}\); but it can easily be ‘acoustically iconic’, \(\text{e.g.}, \text{“hiss”}, \) resembling the sound it refers to. While spatial iconicity is more backgrounded in speech and writing than in gesture, they are also organised according to concepts of space. That even the structure of more arbitrary sign systems is structured by spatial criteria motivates searching for spatial structures in gesture form, since here, spatial iconicity is undoubtedly a cardinal semiotic relation.

The preceding chapters argued in favour of the existence of specific forms in how we conceptualise gestures. This section discusses different general kinds of form in spoken and written language. In Cognitive Linguistics, the conceptualisation of language at the level of cognitive primitives is an established field of study as demonstrated by the lifework of Talmy, Levinson and Langacker.

Talmy’s (2000:191) systematic account of conceptualisations across different prepositions provides us with a distinction along different kinds of forms. He examined the geometric forms that are acceptable in a simple figure-ground construction such as ‘[figure] is near [ground]’. Although a wide range of specific forms are acceptable, Talmy finds that the ground in a ‘near’ construction is conceptualised as belonging to a specific category of form. The ground is

\(^{27}\) Gesture form primitives are in line with Lakoff’s understanding of “image schemas” and “primitives”: “spatial-relations concepts can be decomposed into universal cognitive primitives that recur across languages. […] We call those primitives image schemas” (Lakoff 2006:153-4).

\(^{28}\) Emoticons may come closest to the spatial iconicity in gesture. For example “:-P” coarsely mapping the rough form of punctuation marks onto the eyes, nose, mouth and tongue. See also Section 7.4.
conceptualised as a zero-dimensional form, a point. By varying ground dimensions, he identifies prepositions to require specific ground geometries, as illustrated in Figure 5.1.

<table>
<thead>
<tr>
<th>Different prepositions</th>
<th>Dimensions of the ground</th>
</tr>
</thead>
<tbody>
<tr>
<td>There was oil ...</td>
<td>... near the woodshed.</td>
</tr>
<tr>
<td></td>
<td>Point-like (0D)</td>
</tr>
<tr>
<td>... all along the ledge.</td>
<td>Linear (1D)</td>
</tr>
<tr>
<td>... all over the table.</td>
<td>Planar (2D)</td>
</tr>
<tr>
<td>... throughout the aquarium.</td>
<td>Volumar (3D)</td>
</tr>
</tbody>
</table>

Figure 5.1: Dimensions of the ground in a figure-ground construction. On the basis of Talmy (2000:191)

What does Talmy mean by 'linear' or 'one-dimensional'? Why is the sentence 'There lay berries all along the curvy way through the woods' also acceptable? Why does a non-straight path like 'curvy way' still belong to the one-dimensional category?

Talmy (2000:25-13, 223-225) adheres to a topological understanding of dimensions, in which one-dimensionality includes any path structure of the 'way' in this example; in mathematical terms, any curve. Topological geometry, itself an area of mathematics and an extension of Euclidean geometry, tolerates bending and curving, in general. Whereas in a common-sense understanding of dimensions a straight line is acknowledged to be one-dimensional, people might categorise a bent line, especially when it is curved so much as to form a circle, as being two-dimensional. This is because we need coordinates on at least two axes to define the position of all the points on that line, in a Euclidean understanding. But topological dimensions refer more to the space that the bent line itself occupies. By bending a line, the line does not become a plane; it does not become two-dimensional. Rather, it occupies as much space as before, it still is one-dimensional. A topological understanding of dimensions accounts for the bent line still being a one-dimensional line. Euclidean geometry is thus not applicable to describe the way we conceptualise forms in language and cognition, since the concept of a bent line, in language and cognition, has to belong to the category of lines, not to the category of planes and surfaces.

In topology—and following Talmy, this applies not only to the conceptualisations of figure and ground of spatial arrangements, but to schemas in language in general (Talmy 2000:25-31)—a straight line, a curved line, even a curved line coming to a full circle, are seen as one-dimensional. Equivalently, any kind of surface is two-dimensional. This includes straight
planes as well as baroquely curved surfaces, even a surface closed to a hollow sphere, see Figure 5.2. Paths and surfaces are each regarded as a general schema that spans different specific forms. According to Talmy, a schema “abstracts away from any specificity as to shape (curvature) or magnitude for these points, lines, and planes” (2000:223). Dimensions are understood as the intrinsic space of an object and specify whether a form can accommodate a (bent) 0-, 1-, 2-, or 3-axial coordinate system (Sowa 2006). In a mathematical explanation, any path structure is parameterisable as a one-dimensional line and any surface is parameterisable as a two-dimensional plane.

![Figure 5.2: Forms of one and two dimensions in topological distinction.](image)

It is to be tested whether gesture form can be categorised in dimensions and whether these dimensions are topological rather than Euclidean—just as Talmy found for prepositions in language. A systematic parallelism between language grammar structuring (imaginary) space and gesture form structuring (imaginary) space would be an overlap in a fundamental design feature of language and gesture.

5.2. Profiling

A part of what Langacker investigates as construal of conceptual content is different degrees of “prominence”. One of the operations which focus the attention on one part of content is “profiling” (Langacker 2008:66 et seqq.). As in Talmy’s figure-ground relations above, profiling is the capacity of a linguistic expression to strongly focus on one structural aspect of a concept. The rest of the concept, the “immediate or maximal scope” (ibid.), is not neglected. They are the necessary framework in which the profiled part is to be understood. Langacker’s initial example is that of one body part, for example the hand, being profiled within the larger scope of the whole body (ibid. and p. 64; see Figure 5.3).
Profiling is applied in Sign Linguistics and gesture studies to describe how certain parts of larger concepts are foregrounded by bodily action (Andrén 2010:306; McCleary & Viotti 2010; Mittelberg 2013; Williams 2008:216-26) and how multimodal expressions afford profiling certain aspects of a concept (Parrill & Sweetser 2004; Sweetser 2009), or certain aspects of a metaphorical concept (Cienki & Müller 2008; Müller 2010; Müller & Tag 2012). Gesture Form Analysis will use the word ‘profiling’ to specifically refer to relations within a particular complex material anchor (Hutchins 2005): the gesturer’s body with all its articulators. When a specific part of the body, for example a hand, is profiled in a specific gesture, the immediate and maximal scope of the whole body remains a necessary framework for interpretation. In extending the flat hand, palm up, to present a discursive element to the interlocutor, the hand is profiled as the articulator in contact with the imaginary object and thus more in focus than other parts of the body (Figure 8.29). Still, the whole body is the framework within which the action of the hand is to be understood. The hand extends away from the gesturer’s body allowing visual access to the object and moving it toward the interlocutor. The hand’s position in relation to the whole body is a relevant aspect in the interpretation of this gesture. It is the larger scope within the action of the profiled hand is to be understood.

I will use the term ‘profiling’ to propose operations of gesture form conceptualisation that share the focussing on one part of an entity as in Langacker’s use of profiling. In the relation between an entire unit and a focussed part, profiling overlaps with the notion of a “totum pro parte relationship” (Dirven 2003:94), “WHOLE FOR PART metonymy” (Brdar 2009:263), or “internal metonymy” (Mittelberg & Waugh 2009:343; see Section 2.4). While Talmyn’s use of topology is a general structural aspect of Gesture Form Analysis, profiling specific body parts (articulators) is a process found specifically in the first operations of gesture conceptualisation (Chapter 7).
6. Forms in signed languages

In addition to the Sign Linguistics work cited above, I would like to touch upon work that deals with the general relation of gesture and sign and the way signers use their articulators to represent different entities. These findings are relevant for gesture analysis, too, and will be premises when elaborating the theory Gesture Form Analysis in Chapters 7 and 8.

6.1. Sign is part of gesture is part of action

In the literature on gesture as well as signed languages, we find the use of central terms in a hierarchy that states clear affiliation (Wilcox 2004, Kendon 2004). Every sign is a gesture and every gesture is an action (Figure 6.1).

This hierarchy is not a statement on the importance of the three phenomena, but on their specificity. “To posit a link between gesture and signed languages [...] does not deny that each is unique” (Wilcox 2004:135). Gesture is a specific kind of action in that it is deliberately communicative. A sign language sign is a specific kind of action and also a specific kind of gesture in that it is conventionalised and part of a full-blown language. Regarding the relation of gesture and sign, Kendon supports the relation illustrated in Figure 6.1. He sees “‘gesture’ and ‘signs’ on common ground” (Kendon 2004:307). Schembri, Jones & Burnham also acknowledge the common ground of the three phenomena:

natural signed languages may share some properties with gesture, especially in the meaningful use of space to indicate locative relationships between referents and to indicate participants involved in actions ...

Schembri, Jones & Burnham 2005:272
Nevertheless, each phenomenon is profoundly different. Signs overlap with gestures for example in that both are used to communicate, yet signing is not similar to gesture in many other respects. Psychological studies provide evidence for the differences between the neural processing of so-called “linguistic and nonlinguistic gestural systems”, by which the difference between a sign language sign and pantomime is meant (Corina et al. 1992). Gesture and sign diverge widely in the aspects of conventionalisation, with regard to grammar and lexicon. Lastly, the role of gesture as part of both the natural use of spoken language and sign language should not be underestimated.

The diagram in Figure 6.1 is relevant for this dissertation because the question of how visual action creates form is a question concerning the groundwork of the gestural modality per se (including sign language). It is equally relevant for Gesture studies and Sign Linguistics. It addresses the question of how manual action is perceived to map onto events in the physical world and abstract concepts; in other words, of how gestural action (and signs) can be iconic. As cited from Kendon and Streeck, the word iconic rather is an umbrella term for a heterogeneous group of interpretational strategies that involve similarity in any way. Chapter 7 will attempt to break iconicity down to underlying operations of gesture conceptualisation.

6.2. Blends in depicting scenes

Looking at how articulators are able to represent entities or actions, Sign Linguistics offers an advanced analysis of body actions used for communication in terms of different strategies to create form—as already shown at the example of Mandel’s sketching (Section 2.5). Different approaches in gestures studies adapt notions from the framework of blending (Müller 2009; Parrill & Sweetser 2004; Sweetser 2009), of which some will also be applied in this dissertation.

Liddell (2003) applies Fauconnier’s and Turner’s (e.g., 2002) blending theory to describe so-called real-space blends in sign language use. I rely on Liddell’s definitions of “real space”, by which our concept of the physical world around us is meant, and “real space blend”, which concerns the conceptual integration of real space with another conceptual space. For example the integration of the real space concept of index and middle finger wiggling alternately and

---

29 'Conceptual blending’ is Fauconnier’s and Turner’s explanation of The way we think (2002)—the cognitive operation of creating a new concept from two existing concepts. “The essence of the operation is to construct a partial match between two input mental spaces, to project selectively from those inputs into a novel ‘blended’ mental space, which then dynamically develops emergent structure. Mental spaces are small conceptual packets constructed as we think and talk, for purposes of local understanding and action—they are very partial assemblies containing elements, structured by frames and cognitive models” (Fauconnier & Turner 2003:57-8).
the imaginary space of a walking person may blend into one integrated conceptual space in which the fingers are legs (for an astute treatise of how real-world objects are “surrogates”, consult Liddell 2003:141-75). In these blends, each of the signer’s articulators can stand for a variety of things including persons, specific body parts or other objects. In this way, a signer can even depict multiple entities in complex scenes, all at once. For example in Liddell (2003:142-75), a signer describes a scene of a football coach berating his team (reproduced in Figure 6.2). The signer’s fingers represent the football players sitting in a row (a), while the straight index finger of the right hand represents the coach pacing back and forth in front of them (b). The signer’s face alternately represents the players’ sheepish (a) and the coach’s angry (b) facial expressions. In Dudis’ (2004) terms, the signer’s body is “partitioned”.

![Figure 6.2: (a) A |football team| sitting down represented by curled fingers being berated by (b) their |coach| represented by the index finger of the right hand. From Liddell (2003:310; excerpt). Courtesy of Cambridge University Press. Copyright © 2003, Cambridge University Press.](image)

Every part of the body can be involved in a separate kind of representation, even simultaneously. In an articulator standing for something else, the sign linguists Liddell and Dudis both use straight brackets to talk about the imagined object the articulator stands for. I will adopt these authors’ convention for naming body parts which are being interpreted as standing for something else, but use it, more broadly, even when an articulator stands for something that is not of the same size (e.g., Figure 6.2). In the telephone gesture (hand between ear and mouth, thumb and pinkie extended), the hand stands for the |telephone|, in the football coach example (Figure 6.2), the fingers stand for the |players| and the |coach|, while the face stands for the |players’ faces| or the |coach’s face|. Also objects that are to be imagined outside of the space occupied by the articulators are notated as |objects| (e.g., Figure 2.7).

It remains open whether other gesture types, such as Mandel’s sketching (using the index finger to draw an imaginary line in the air or other examples of trace leaving) are also
combinable with the different kinds of body partitioning. At least for gesture, this seems to be the case. The hand can point to another articulator that represents an object (Figure 8.45) or the right hand may draw on a specific shape represented by the left hand (Figure 8.22). It is within this blend, which integrates real space with an imaginary space, that gesture and context evoke meaning, in which a specific gesture form, whose structure will be now proposed, is created.
Part II:

Gesture Form Analysis as basis for a new gesture typology
Part II presents theory and application of Gesture Form Analysis. This model allows for a formal description of the spatial structures encountered in gesture analyses in gesture studies and related areas, as reviewed in Part I. Gesture form is structured according to two principles: a) gesture form can be described in terms of a set of four topological form primitives; b) gesture form is elicited by a combinatorial series of eight geometric operations.

The primitives of gesture form posited by the first principle are dimensions in a topological understanding, as analysed in Talmy (2000) with regard to language on space. Any form is conceptualised as zero-, one-, two-, or three-dimensional. The second principle is constituted by gesture form operations. These spatial operations organise how a shape aspect is abstracted from the shape of the articulator and how derived forms can be computed according to geometric rules. Introducing (Chapter 7) and exemplifying (Chapter 8) these operations is an important part of the attempted contribution to gesture studies. The gesture examples will encompass gestures from my Studies 1 and 2, which were conducted for this dissertation, as well as diverse examples from the literature. The proposed method for describing gesture form in detail, Gesture Form Analysis, will apply the proposed eight operations in a modular way. It will include a graphical illustration of the coding scheme that displays the forms that an observer has to process in order to understand, for example, a pointing or a moulding gesture. The analysis of several different gesture types will have, as additional product, a typology that, in contrast to other functional gesture categorisations, is strictly based on gesture form. Every class in this typology, every gesture type, will be defined in all its necessary spatial operations and marked as such by upper case initials (e.g., Pointing At Object).

7. Two principles of gesture form

This chapter presents a model of gesture form. An investigation of form that systematically distinguishes between the physical form of the articulators and the sequence of imaginary forms required for understanding the gesture, has not yet been carried out. This chapter will focus on proposing the principles of such a model of gesture form. The combinations of gesture form operations will provide with a matrix of all theoretically possible gesture types. Chapter 8 will then analyse the combinations found in empirical data, building up to a basic

---

30 Study 1 will be discussed in Chapter 8; Study 2 will be dealt with in detail in Chapters 9 through 11 (Part III).
gesture typology. I would like to apply the following proposal by Levinson and Haviland's about linguistic modelling to gesture. They advocate the conviction that, we “do indeed need to construct a serious semantic typology, and search for underlying patterns and uniformities on a quite abstract level” (Levinson & Wilkins 2006:514).

For a systematic description of gesture form, this dissertation builds on two distinctions found in the literature, which were discussed in part I: the insight that, as part of language structure, objects are conceptualised as being of zero to three dimensions (Talmy 2000), and the perspective on gesture of the hand being conceptualised as different tools (Müller 1998; 2010; 2014; Streeck 2009). These two perspectives supply the model of gesture form with one principle each. The first supplies form-dimensional primitives, while the second, in its search for underlying structures of the hand as different tools, will yield the operations that elicit and specify the constituent forms that make up gesture form.

The literature review in Chapter 3 highlighted forms of certain dimensions occurring in specific gesture types. Selected gesture types were analysed in two groups. Pointing and drawing gestures on one hand, and surface-contact gestures on the other. Gesture types within these groups were structured in preliminary hierarchies, depending on the forms consistently recurring in their analyses. These preliminary hierarchies allow for first predictions on how gesture form could account for the relations between gesture types. The results are summarised in Figure 7.1.

This chapter will propose operations that underlie the group of pointing and drawing gestures, and the group of surface contact gestures, as well as kindred gesture types within them.
Together with further gestures that call for a separate group, they will be integrated in a single model. In this model, the conceptualisation of the articulator as a shape of one, two or three dimensions is the basic distinction (‘Shape Profiling’, Section 7.2.3). For example, comprising pointing and drawing, the articulator has to be profiled as a one-dimensional line. Otherwise, indicating a referent or drawing a shape in the air would not be possible. In surface-contact gestures, the two-dimensional (tactile) surface of the articulator is profiled, since this surface is to be imagined in contact with the referent’s surface. The gesture type called representing stands apart in capturing the whole three-dimensional form of the articulator—the hand can stand for the referent not only in (part of) its surface, but in its three-dimensional volume—while also showing its movement.

Figure 7.2 illustrates this very broad distinction, thus accounting for the difference in the profiled dimensions between gesture types. This structure parallels the final gesture typology in that all gesture types (right column) are grouped by the dimensions of the profiled shape (left column). However, these distinctions are rather crude and also non-complete with regard to the model presented in this chapter, but they are intended to attune the reader to the kind of typology this and the following chapter will build up to.

<table>
<thead>
<tr>
<th>Profile</th>
<th>Gesture type</th>
</tr>
</thead>
<tbody>
<tr>
<td>1D</td>
<td>pointing (directing)</td>
</tr>
<tr>
<td></td>
<td>pointing (at something)</td>
</tr>
<tr>
<td></td>
<td>drawing</td>
</tr>
<tr>
<td>2D</td>
<td>holding</td>
</tr>
<tr>
<td></td>
<td>moulding</td>
</tr>
<tr>
<td>3D</td>
<td>representing in 3D</td>
</tr>
</tbody>
</table>

Figure 7.2: Simplified distinction of gesture types according to the profiled shape.

Contrasting to related literature in gesture studies, the technical terms of Gesture Form Analysis used in this dissertation will not include iconicity, metaphor, metonymy, directionality, and indexicality. This is not because these terms deal with something different, but rather because I would like to promote a set of operations that adapt the semiotic relations captured by these terms specifically to the spatial reality of gesture.

This dissertation’s main aim being to redefine gesture types with the help of a model of primitives and operations, the following two sections will theoretically introduce both of these principles. Section 7.1 defines spatial dimensions in gesture, challenging the use of dimensions
in accounts of Müller (1998) and Sowa (2006). I will argue that only in a topological understanding of dimensions (Talmy 2000) can dimensions serve consistently as primitives of form, given the detail of structural analysis is to be at the level of gesture form. Section 7.2 introduces the operations of Gesture Form Analysis, which will provide the structural elements of gesture typology. Section 7.3 compresses the proposed structure of gesture form into a definition, whereas Section 7.4 regards the proposed geometric operations in the light of semiotic relations widely used in gesture studies.

Making claims about how the neurological processes constituting the conceptualisation of gesture form is clearly beyond the scope of this dissertation. As will be illustrated empirically in Chapter 8, the dimensions and operations laid out in the following serve to describe the simple geometrical procedures necessary for understanding certain gesture types. How these geometrical operations are computed in the brain cannot possibly be investigated with the methods at hand (introspection, gesture analysis, motion-capture data, or theory of affordances). This would require (neuro-)psychological studies. The claim of Gesture Form Analysis is limited to that, in one way or another, certain geometrical operations have to be completed in order to understand a specific gesture (see, e.g., Section 8.3.2). For example the conceptualisation of an index-finger-pointing gesture targeting a car across the street necessarily contains 'linear continuation' of the extended index finger. I think it is uncontroversial that, without this continuation, the unambiguous identification of a distant object could not be explained. As will be argued by the analyses in Chapter 8, the spatial operation of continuation is a 'necessary condition' for this gesture type. By this approach, I follow Goffman’s guiding “question as to how interaction is possible in the first place” (Kendon 1988:19). The gestural function of showing a specific car to an interlocutor requires the conceptual process of continuation. Unanswered remains the specific question of how the brain accomplishes 'linear continuation' or how this computation overlaps for example with processes of visual perception or language interpretation. In short, the proposed analysis is concerned with whether certain spatial operations take place, not how.

7.1. Dimensions: Gesture form primitives

Chapter 3 documented gesture scholars’ use of terms such as ‘pointing’ or ‘moulding’ when speaking of gestures that appear to not be equivalent in gesture form. How can we systematise the description of these forms? Are there markedly different categories of form? In other words: are there form primitives or at least ways to exactly describe the different forms we see in a pointing gesture compared to, say, those in a moulding gestures?
7.1.1. Spatial dimensions in gesture

In general, a gesture can elicit forms of different dimensions, which are the basis of meaning interpretation. Gestures can mark a point in space, show a direction, mould the surface of something or represent a physical object. Figure 7.3 exemplifies forms of zero to three dimensions that can be elicited by certain hand shapes.

![Figure 7.3: Four different types of form in gestures conceptualisation.](image)

How are these forms related to different gesture types? What role do they play in fulfilling different gestural functions? The following section describes characteristics of different forms with regard to what they afford. This is to lay the grounds for distinguishing different kinds of forms, that is, forms of different (topological) dimensions.

7.1.2. Different form dimensions afford different states and actions in space

As derived from Gibson's (e.g., 1977) work on affordances of surfaces (Chapter 4), this section treats affordances of the whole spectrum of zero- to three-dimensional forms in one- to three-dimensional space. The proposed gesture typology will adhere to these affordances. In light of their special affordances, forms of different dimensions (form primitives) are one principle of a model of gesture form. This and the following sections are directed towards form dimensions as specific means of expression, as the most basic distinction of gesture form.

While the primitives of gesture form are topological dimensions (as will be discussed in this chapter), the dimensions of gesture space are Euclidean. Gesture form includes curved paths that belong to the category of lines (topological categorisation), but the space in which gestures are executed in is not curved but organised by straight axes (x, y and z). This will become specifically important in Section 7.2.11.

Gibson scrutinised the affordances of surfaces. The emphasis on surfaces in Chapter 4 within the discussions of gesture perception is to subscribe to his claim that surfaces play a key
role in identifying the affordances of objects around us (Gibson 1977). The frequent occurrence of Holding and Moulding (Section 8.5) illustrates the importance of form interpretations involving surfaces, in gesture. An example for one of these surfaces is the ground that we walk on. “The ground is literally a basis for behavior, and also a sort of basis for visual perception” (ibid., p. 72). Since we are especially interested in form, how exactly is a surface the literal “basis for behaviour”? Gibson says it “is lie-on-able and stand-on-able [... it] affords locomotion” and that surfaces also afford “concealment” (pp. 72-73). What do these two affordances, being “stand-on-able” and “concealment”, have in common? In both cases, the surface separates the substance on one side of it from the substance on the other side of it. Our feet do not merge with the ground when they push against it. We can stand on the ground, because the surface of the solid ground affords that our feet can stem us up against gravity. Furthermore, we can walk on the ground because the surface of the ground also affords exerting lateral force onto it to move our body forward in a desired direction.

The surface of a solid object also defines strictly where this object ends and something else begins. We can push or grasp a rigid object because its surface affords using bodily movement and force (in combination with traction) to push and control the object. If the object is opaque, it also means that our visual perception of the object will be constraint to everything on that surface, making its whole interior visually inaccessible.

In three-dimensional space, the surface is the most efficient border imaginable, because a surface has a thickness of zero. It does not occupy any volumetric space. Still, it divides portions of space, volumes, like a plastic wrap. A plastic wrap has very little mass but still divides two substances from one another, thus its function of isolating food from bacteria or other substances from the outside. Just as a plastic wrap, a surface can be closed to a sphere or around any other three-dimensional form. In this specific case, division becomes absolute; the affordance of division becomes the affordance of enclosure. Now, the surface encloses a volume of space. This kind of enclosure is what we see in any real-world object: objects are enclosed by their surfaces. The affordance of enclosure is present in Talmy’s work (2000:191-5) and, under the term “concealment”, also in Gibson’s (1977:73-4).

Is division a general affordance of two-dimensional surfaces? If so, can we find the operation of division in the way we conceptualise gestures, too? Are there specific and distinct affordances for other dimensions, the point (0D), the line (1D) and the volume (3D)? Gibson’s work on the affordances of surfaces leaves open whether other dimensions of form also afford specific special states, and consequently specific behavioural strategies. This chapter will
attribute affordances to forms of other dimensions. This basic form-function mapping suggests form dimensions as viable candidates for gesture form primitives.

Chapter 3 identified a line and sometimes a point as components of understanding pointing. As pointing gestures function, for example, to direct, single out (imaginary) objects, and disambiguate potential referents, there is a range of possible affordances that could be attributed to lines and points. These general functions of gesture types will therefore be broken down to the form dimensions of line (1D) and point (0D). The result is an attempt to consolidate the theory of affordance and gesture conceptualisation, whose analytic fruit, Gesture Form Analysis, will then be variously applied in Chapter 8.

Affordances of a point (0D)

A zero-dimensional shape, the point, is *per se* the simplest shape, or strictly speaking, it is no shape at all, just a location. No matter whether in one-, two- or three-dimensional space, the central affordance of a point for communicative purposes is that it can be used to precisely indicate a specific location. This affordance is essential in the gesture types that point at something. In two- and three-dimensional space, this seems to be the only affordance of a point.

Figure 7.4 further illustrates the point’s affordances of division and enclosure in one-dimensional space. In one-dimensional space, a single point affords dividing such a space into two (a). Two points additionally afford enclosing the space between them: a line segment (b). The affordance of division can be elaborated, by recurrence, to enclosure.

![Figure 7.4: One point (a) divides 1D space, two points (b) enclose 1D space.](image)

Affordances of a line (1D)

Considering a line and higher-dimensional forms, it is necessary to distinguish whether the form is straight or curved. A straight line provides with a direction, indispensable in pointing and drawing gestures. It is also the shortest path from one point to another, at least in Euclidean geometry. Therefore, the direct path from one extreme of a shape to the other is also a simple measure of size. A curved line, in addition to projecting a directional curve, can also be used to describe a shape contour as well as the form of a movement trajectory.
In two-dimensional space, a single straight line can afford dividing a two-dimensional area into two (Figure 7.5a), while two straight lines (b) or one curved line (c) can also enclose an area partially (b) or completely (c). The different specifications of a referent in partial and complete enclosure will be discussed later in this chapter (Section 7.2.8), in the following chapter (Figure 8.30 through Figure 8.43) and in the analyses for Study 2 (Section 10.3).

Figure 7.5: One straight line (a) divides 2D space, two (or more) straight lines (b) or one curved line (c) enclose 2D space.

Of course, lines can also divide and enclose one-dimensional space, but since points afford the same state more precisely and by a simpler, i.e. lower-dimensional, form, these less efficient affordances are backgrounded in this theoretical presentation. The same is true for two- and three-dimensional forms (see also Figure 7.7).

Affordances of a surface (2D)

A single straight or curved plane divides three-dimensional space (Figure 7.6a). Two planes (b) or a curved surface (c) can also enclose three-dimensional space partially (b) or completely (c). This affordance is crucial in our three-dimensional world, which consists of three-dimensional entities only. It is also the basis for the very large group of gestures that touch, hold or otherwise interact with objects (Section 8.5).

Figure 7.6: One plane (a) divides 3D space, two (or more) planes (b) or one curved surface (c) encloses 3D space.

Gesturing hands can convey the shape of an object, if they are able to depict surface properties of this object. Gestures can convey imaginary objects (henceforth |objects|), by the interaction of the hands with the object’s surfaces. Showing actions on |objects| specifies surface
characteristics of these objects by the way the hand is shaped in relation to the performed movement. Inferring the object’s shape specifically relies on profiling the hands’ surfaces.

**Affordances of a volume (3D)**

Three-dimensional shapes instantaneously afford representing complex shape information (Müller 2010; 2014). They can represent objects not only in their shape, as discussed in Section 3.5, but, when moved, they additionally allow giving a complex movement description, including rotation and changes of shape including bending, squeezing and stretching.

### 7.1.3. Summary of how different forms afford division

At this stage, it will be summarised which forms afford division in one- to three-dimensional space (this section) and then what the most prominent affordances of zero- to three-dimensional forms are in three-dimensional space (Section 7.1.4). The affordances illustrated in Figure 7.4 through Figure 7.6 will be referred to in these short summaries and underlie not only the remainder of this chapter, but also all analyses in Chapter 8.

The fact that generally, in a space of x dimensions, a form of x-1 dimensions can afford dividing space is presented by Figure 7.7 in tabular form. This is relevant to gesture, since many gestures are conceptualised in low-dimensional space (like on a map or a projected profile view (2D), see, e.g., Figure 8.9, Figure 8.42 and Figure 8.43, or just showing the (1D) extent of an object, see, e.g., Figure 8.30 through Figure 8.32).

<table>
<thead>
<tr>
<th>Does a form afford division?</th>
<th>0D form</th>
<th>1D form</th>
<th>2D form</th>
<th>3D form</th>
</tr>
</thead>
<tbody>
<tr>
<td>1D space</td>
<td>Yes</td>
<td>Yes</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2D space</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
<td></td>
</tr>
<tr>
<td>3D space</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
</tr>
</tbody>
</table>

Figure 7.7: Forms of 0-3D affording division across one-to-three-dimensional space. The simplest forms (lowest number of dimensions) that afford enclosure are in bold face.

As introduced in Section 7.1.2, enclosure is a specification of division (both being a specification of adjacency, see Section 7.2.6). When forms of specific dimensions can divide space, they can also enclose it. Forms can divide space partially, by using for example two shapes to partially enclose the in-between space, or completely, by a single curved shape (or at
least three straight or plane shapes), see Figure 7.4 through Figure 7.6. In Chapter 8, dividing space and especially enclosing space will be illustrated as a powerful tool to depict the shape of an object.

Lastly, a form of specific dimensions generally affords depicting another form of the same dimensions. It is self-evident but not irrelevant that a form can in itself represent another form of the same dimensions. It can also characterise higher-dimensional forms by portraying aspects that are of the same dimension as itself, or aspects that are conceptualised as being of the same number of dimensions.

7.1.4. Summary of specific affordances in three-dimensional space

Figure 7.8 offers an overview of affordances of different forms in (three-dimensional) gesture space, as discussed in this section up to now. I suggest that these affordances are crucial for pointing gestures (0D), direction giving (1D), interaction with objects (2D) and ‘body part as object’ gestures (3D).

<table>
<thead>
<tr>
<th></th>
<th>0D forms</th>
<th>1D forms</th>
<th>2D forms</th>
<th>3D forms</th>
</tr>
</thead>
<tbody>
<tr>
<td>Location</td>
<td>Location</td>
<td>Direction</td>
<td>Enclosure</td>
<td>Representation</td>
</tr>
</tbody>
</table>

Figure 7.8: Prominent affordances of form dimensions in three-dimensional space.

By listing conceptual and communicative affordances of forms of different dimensions, I would like to propose form-function mappings as the foundation of gesture analysis. Figure 7.8 focusses on only one the most prominent affordance per dimension. In addition to there being more than one affordance per form dimension, only the combination of different forms (of 0-3D) can fulfil the functions associated with certain gesture types analysed in Chapter 8. Many pointing types, for example, include a one-dimensional form affording indicating a direction as well as a zero-dimensional form indicating a location.

7.1.5. Fine-grained analysis for distinguishing spatial dimensions in gestures

I would like to claim that several forms are part of conceptualising one gesture. So far only few studies have implicitly argued for the existence of a combination of multiple forms in gesture conceptualisation (Mandel 1977; Sowa 2006). The systematic interactions found in these approaches are the starting point for a new gesture typology.
Gestures can elicit forms of different dimensions. Gestures can mark a point in space, show a direction, mould the surface of something or represent a physical object. Figure 7.9 displays forms of zero to three dimensions that can be elicited by specific hand shapes.

![Figure 7.9(a-d): Forms of different spatial dimensions in gestures.](image)

The hand shape with which we draw attention to a point in space (Figure 7.9a) can be the same as the hand shape that indicates a direction (b). The fact that one shape can elicit two distinct forms opposes the idea of one-to-one mapping of hand shape (with or without movement) and elicited form. This dissertation proposes discerning two things: (1) the physical articulator form and the forms that are part of an observer’s conceptualisation of articulator form: gesture form. (2) Within one gesture, multiple alternative forms are considered as well as the possibility of multiple constituent forms making up gesture form. Forms of zero to three dimensions, as exemplified in Figure 7.9, are the primitives of gesture form.

The conceptualisation of a stretched out articulator (such as the index finger of a pointing hand) may include a line that extends in a specific direction (Figure 7.9b). To arrive at a conceptualisation that includes a point at the tip of the finger (Figure 7.9a), the gesture form will be claimed to involve two imaginary forms. One is a line the finger is conceptualised as, and the other is the point of intersection of this line with the surface of the hand at the tip of the finger. Thus, figures (a) and (b) both rely on conceptualising the articulator as a line, while (b) continues the line to stand for a (one-dimensional) direction and (a) focusses on the intersection of this line with the surface at the tip of the finger in a point.

The comparison of Directing versus Pointing At Location (Sections 8.4.2 and 8.4.3) introduces the idea of different combinations of constituent forms in gesture, since both types can be performed with the same hand shape. Figure 7.9 does not specify operations that could account for different interpretations of the same hand shape. The following section will do so,
suggesting that a set of operations, as part of a model of gesture form, are applicable across types as different as pointing, drawing and moulding.

7.2. Gesture form operations

The distinction of forms according to the three possible kinds of extensions a form can have (1-3D) plus the non-extending point (0D) is implicit in Mandel’s research on iconicity in sign language (1977) and in current computational approaches to gesture (Kopp 2005; Martell 2005; Martell et al. 2002; Rieser 2010; Sowa 2006). These categories are taken to be primitives of gesture form. This section will propose gesture form operations as the junctions between the different forms of a single gesture. There are eight gesture form operations. Their capitalisation is to mark them as formalised geometric operations introduced in the following sections: Articulator Profiling and Shape Profiling (Section 7.2.3), Continuation (7.2.4), Intersection (7.2.5), Adjacency (general) (7.2.6), Adjacency (Division) (7.2.7), Adjacency (Enclosure) (7.2.8) and Trace Leaving (7.2.9).

The gesturer’s body is the starting point of gesture form. Not only the hands, but any part of the body can produce a gesture—alone or combined into “articulator units” (Hassemer 2009; Section 3.5.1 on articulator independency). The form of the body, \( b \), and the form of the profiled articulator, the articulator form, \( af \), have in common that they are members of the set of all three-dimensional forms:

\[
\begin{align*}
(1) & \quad b_{3D} \in \{ z \mid z \text{ is a } 3D \text{ form} \}. \\
(2) & \quad b_{3D} \in 3D \text{ forms} \\
(3) & \quad af_{3D} \in 3D \text{ forms}.
\end{align*}
\]

Henceforth this will be abbreviated as:

\[
(4) \quad x \in 0D \lor 1D \lor 2D \lor 3D \text{ forms} \tag{31}
\]

An operation, \( p \), is a function, or, more generally, a mapping, acting on an input form \( x \). An input form can be of any dimensions.

\[
(4) \quad x \in 0D \lor 1D \lor 2D \lor 3D \text{ forms} \tag{31}
\]

\(^{31} \lor \) stands for a logical ‘or’.
In what way an operation modifies the input form has to be specified for each gesture, but every operation changes the dimensions of an input form according to a set of rules, producing an output form $u$ of certain dimensions (see also Figure 7.10). This section will give equations that specify possible dimensions of the input forms and the respective output forms, going through all gesture form operations. An output form can also be of zero, one, two or three dimensions:

\[(5) \quad p(x) = u_{nD}; \quad u_{nD} \in 0D \lor 1D \lor 2D \lor 3D \text{ forms.}\]

In short:

\[(6) \quad p(x) = u_{nD}; \quad 0 \leq n \leq 3.\]

To illustrate this and some of the following formulas, I give equations that illustrate the form specifications in the regular equation by drawings. These graphical equations will carry the label “ANY” when only the dimensions of the shape are relevant instead of the specific shape displayed in the equation. (7) is the graphical form of (6).

\[(7) \quad p(\text{ANY}) = \text{ANY} \lor \text{ANY} \lor \text{ANY} \lor \bullet.\]

The gesture form operations are seen as mappings, that is, $p$ acting on one specific constituent form, the input form $x$. They are concatenated mappings. The output form of one operation is the input form for the next operation. For example, if two operations define the gesture form $gf_2$ of a gesture whose form is captured by input form $x$, then

\[(8) \quad gf_2(x) = p_2(p_1(x)), \quad \text{henceforth simply: } p_2p_1(x).\]
This means, gesture form may be composed of \( n \) operations that are concatenated mappings. The first operation acts on the first constituent form, its input form, \( x \), producing the output form \( p_1(x) \), the second constituent form of the overall gesture form, \( g_f \). The form \( p_1(x) \) is the input form for the second operation, \( p_2 \), which outputs the third constituent form \( p_2p_1(x) \). In a gesture with two operations only, the third constituent form, \( p_2p_1(x) \), would then be the resulting form. In a gesture with more operations, respective constituent forms would follow.

The first input form is the body \( b_{3D} \). Accordingly, the general structure of gesture form is

\[
g_f(b_{3D}) = p_n p_{n-1} \ldots p_2 p_1(b_{3D}).
\]

Each operation adds a form to the forms already elicited by prior operations. This is contrary to a regular reading of equation (9): not only the end result contributes to the gesture form, but—usually to a lesser extend or more backgrounded—all constituent forms. This is to avoid extended equations. Comprehensively, equation (9) would be:

\[
g_f(b_{3D}) = p_n p_{n-1} \ldots p_2 p_1(b_{3D}) + p_{n-1} \ldots p_2 p_1(b_{3D}) + \ldots + p_2 p_1(b_{3D}) + p_1(b_{3D}) + b_{3D}.
\]

One gesture’s conceptualisation encompasses all constituent forms and is thus a combination of different forms, of which each is the processing output of the prior form. Still, as Section 7.2.13 will discuss, one form is usually dominant in gesture perception—what I call ‘resulting form’. The abbreviated style of equations shall emphasise the resulting form. Nevertheless, all constituent forms are part of gesture conceptualisation.

\[7.2.1. \text{Order of operations}\]

Gesture Form Analysis proposes operations occurring in a specific order. Gesture perception of a naïve observer has to follow this order, while observers (including the gesturer) that have prior knowledge about the referent might be able interpret a gesture in a different order of operations, for example, drawing conclusions on what the gesture means from what they already know about the referent.

There are two characteristics of the order of operations as shown in (11) and (12). Operations are non-commutative. The order of concatenated operations is relevant:
(11) \[ p_i p_j(x) \neq p_j p_i(x); i \neq j. \]

Exceptional cases, in which operations are generally interchangeable, will be indicated (e.g., Figure 8.41).

Any operation can occur more than once within the constituent forms of one gesture, as for example in

(12) \[ g f_n(b_{3D}) = p_n p_{n-1} \ldots p_3 p_2 p_1 b_{3D}; \quad \text{with, e.g.,} \ p_2 = p_4. \]

7.2.2. Specific order: profiling operations in the beginning

The first gesture form operation is always the operation of Articulator Profiling, \( p_{AP} \), which is followed by the operation of Shape Profiling, \( p_{SP} \). This specifies the general structure of gesture form given in (9) to:

(13) \[ g f_n(b_{3D}) = p_n p_{n-1} \ldots p_3 p_{SP} p_{AP} b_{3D}; \quad \text{(structure of gesture form).} \]

Figure 7.10 illustrates this general formula, also emphasising the multi-layered character of gesture form: constituent forms are output forms of one operation and input forms for another operation.

![Figure 7.10: Gesture Form Analysis coding scheme.](image-url)
The operations of Articulator and Shape Profiling

Operations can determine the dimensions of their output form, as a function of the dimensions of the input form. This and the following passages (Sections 7.2.3-7.2.9) formalise the constraints of all gesture form operations. Subsequently, Chapter 8 provides numerous examples of practical application of all eight operations.

As an example of how an operation determines the dimensions of the output form, the first operation, Articulator Profiling, \( p_{AP} \), that acts on the three-dimensional form of all potential articulators of the body, \( b \), maintains the dimension of the input form, and outputs also a three-dimensional articulator form \( af \):

\[
(14) \quad p_{AP}(b_{3D}) = af_{3D}.
\]

The operation Articulator Profiling selects one articulator of the range of potential articulators the body offers. Though for example the hand is only a small portion of the whole body, it is, just as the whole body, three-dimensional.

Contrastingly, Shape Profiling, \( p_{SP} \), can change, more precisely reduce, the dimensions of the profiled articulator to two, one or zero dimensions. This is not necessarily the case. It can also maintain the full three-dimensional form of the profiled articulator:

\[
(15) \quad p_{SP}(x_{3D}) = u_{nD}, \quad 0 \leq n \leq 3.
\]

In other words, the articulator can be relevant just in its location and movement (0D), the articulator’s shape may prominently exhibit a linear extend conveying a specific direction (1D), part of the articulator’s tactile surface can be meaningful (2D), or the whole articulator can represent another object in its solid shape (3D).

\[
(16) \quad p_{SP}(\text{[ANY]}) = \text{[ANY]} \quad \text{[ANY]} \quad \text{[ANY]} \quad \text{[ANY]} \quad \text{[ANY]}
\]

The use of ‘profiling’ in Articulator and Shape Profiling is to mark the consonance with the Langacker’s (2008) ‘profile’ within the notion of ‘focusing’, as touched upon in Section 5.2.
To illustrate different forms that Shape Profiling (equation 16) is applicable to, examples of specific input and output forms are given in the graphical equations (17-20):

\[(17) \quad p_{SP}(\square) = \square\]
\[(18) \quad p_{SP}(\square) = \square\]
\[(19) \quad p_{SP}(\|) = \|\]
\[(20) \quad p_{SP}(\cdot) = \cdot\]

7.2.4. The operation of Continuation

The operation of Continuation $p_{CN}$ extends the input form, based on its dimensions and form qualities. This operation applies only to one- and two-dimensional forms. Zero-dimensional points do not extend at all, thus they do not contain information as to how to be continued. Three-dimensional forms extend on all axes, thus they do not specify in which direction they are to be continued. The output form always has the same number of dimensions as the input form. For three-dimensional space,

\[(21) \quad p_{CN}(x_{nD}) = u_{nD}; \quad n = 1, 2.\]

For how long a form is being continued is not determined. A line can be continued in both directions infinitely long. Surfaces, extending on two axes, can be continued infinitely on both axes, thus potentially producing an output form of different proportions than the input form. Examples for straight shapes:

\[(22) \quad p_{CN}(\, ) = \,\]
Examples for curved shapes:

\[(24) \quad p_{\text{CN}}(\bigcirc) = \bigcirc \]
\[(25) \quad p_{\text{CN}}(\bigtriangledown) = \bigtriangledown \]
\[(26) \quad p_{\text{CN}}(\bowtie) = \bowtie \]

7.2.5. The operation of Intersection

The operation of Intersection $p_{\text{IN}}$ reduces the number of dimensions of the input form with regard to another form physically present or imagined to be present by the observer of the gesture. When two forms intersect, the output form has the number of dimensions as defined by geometric intersection. For forms of one and two dimensions, the output form is of one dimension less than the dimensions of the simplest input form (the input form of least dimensions).

\[(27) \quad p_{\text{IN}}(x_{\text{nD}}, y_{\text{mD}}) = x_{\text{nD}} \cap y_{\text{mD}} = u_{(\min(n,m) \cdot 1)\text{D}}; \quad n, m \in \{1,2\},^{33}\]

Examples:

Two non-parallel surfaces (2D) intersect along a line (1D),

\[(28) \quad p_{\text{IN}}(x_{2\text{D}}, y_{2\text{D}}) = x_{2\text{D}} \cap y_{2\text{D}} = u_{1\text{D}}\]

---

$^{33}$ “$\cap$” stands for geometric intersection.
two non-parallel lines (1D) intersect at a point (0D),

\[ P_{IN}(x_{1D}, y_{1D}) = x_{1D} \cap y_{1D} = u_{0D} \]

a line (1D) intersects with a non-parallel surface (2D) at a point (0D),

\[ P_{IN}(x_{1D}, y_{2D}) = x_{1D} \cap y_{2D} = u_{0D} \]

and so forth. Some graphical examples are given in the following (the black shape on the right of the equation is the intersection):

\[ P_{IN}(\text{\textbullet}) = \text{\textbullet}. \]

\[ P_{IN}(\text{\textbullet}) = \text{\textbullet}. \]

\[ P_{IN}(\text{\textbullet}) = \text{\textbullet}. \]

7.2.6. The operation of Adjacency (general)

The operation of Adjacency (as derived from the work of Mittelberg, Jakobson and Peirce, see Section 2.4) does not necessarily constrain the output form’s dimensions. As it deals simply with immediate spatial neighbourhood of two entities, the form of one entity is not necessarily related to the form of the other entity. For example, if we want to point at a table, drawing attention to any point on this table’s surface can be sufficient, as an intermediate step (constituent form), to successfully point at the table as a whole. Any one form can be adjacent to any other form. This is why the very general operation of Adjacency does not constrain the output form at all.

\[ P_{AD}(x_{ad}) = y_{ad}; \quad n, m \in \{0, 1, 2, 3\}. \]
7.2.7. The operation of Adjacency: Division

The operation of Division is more specific than Adjacency in that the \( n \)-dimensional input form \( x \) cuts through the space of a form of one additional dimension \((n+1)\) dividing it into two forms \( u \) and \( v \) adjacent to either side of it. In addition to equation 31, Figure 7.4 through Figure 7.6 already anticipated this and the following operation (Enclosure).

\[
\begin{align*}
(35) \quad p_{DI}(x_{nD}) &= u_{(n+1)D} \lor v_{(n+1)D}; \quad n = 0, 1, 2^{34} \\
\end{align*}
\]

In one-dimensional space, this would be, for example,

\[
(36) \quad p_{DI}(\cdot) = \cdot \cdot \cdot
\]

7.2.8. The operation of Adjacency: Enclosure

Enclosure is yet a more specific kind of Adjacency and also a more specific kind of Division. Thus, as in Division, Enclosure outputs a form of one additional dimension. In contrast to Division, Enclosure focusses on one form, the enclosed form. It usually neglects the other forms outside of the enclosing form.

Enclosure occurs in a wide range of gestures that indicate aspects of size and shape of objects. The two cases observed in the data (Figure 8.30 through Figure 8.36) were an output form (partially) enclosed either by one input form or two input forms of the same number of dimensions:

\[
(37) \quad p_{EN}(x_{nD}, y_{nD}) = u_{(n+1)D}; \quad n = 0, 1, 2.
\]

In Addition to the basic affordances of different form dimensions with regard to Division and Enclosure (Figure 7.4–Figure 7.6), the following graphical equations exemplify some of the many ways in which an output form can be enclosed by one or two input forms.

\[34 \quad \lor \] stands for addition. Two output forms are being produced.
In Enclosure, some of the dimensions of the output form might not be specified. In holding an object between (the surfaces) of the flat hands held vertically, the height and depth of the object are not specified because the object is only enclosed on the left and right side but not at the top, bottom, front or back. The object can extend variably into these directions, as in spoken language, too, calling it “extendability in ungoverned directions” (Talmy 2000:216-7; 2003). While the resulting form can extend into these directions, any extension in these directions can also be backgrounded or completely ignored, as in accompanying the same gesture with ‘the thread was that long,’ in which gesture space can be reduced to display extent on one axis only (Section 7.2.12).

Still, even partial Enclosure can convey some of the object’s shape aspects such as the extent on one axis and parts of the surface curvature. Since the articulator has to be imagined...
continuously adjacent to the |object|, the inner surface of the articulator has often to be understood as the corresponding and adjacent outer shape of the |object| as in equations (39-44). Two articulators can show the |object|’s extent that is to be imagined going from one articulator to the other as in (38, 39, 40, 42, 43). Finally, the articulators can also convey curvature as in (40, 41, 43, 44). The more the articulators enclose the |object|, the more they reveal of the object’s shape. While there are numerous ways to partially enclose an object, it must suffice here to state that in the case of (39 and 42) two straight shapes can specify the output form on one axis—regardless of the complete output potentially being of more dimensions. Regardless of the output form being two-dimensional (39) or three-dimensional for (42), it is specified on one axis only (in (39) and (42) the horizontal axis, see Figure 8.32 for an empirical example).

7.2.9. The operation of Trace Leaving

Gestural motion can be taken to stand for the actual motion of an entity. In these cases, the aspect of motion as part of the gesture phenomenon is conceptualised as what it actually is: the simple change of location of an entity as a function of time. Alternatively, motion can also stand for something static, the extent of an entity. In Talmy’s theory of fictive motion of language (2000), motion stands for extension. Motion also stands for extension when we perceive an index-finger-pointing hand, moved laterally, to draw a line in the air that extends through space. In gesture, a moving fingertip can be perceived to leave a trace.

Mandel does not only imply the functioning of Trace Leaving in its entirety, but also applies this geometric relation to different kinds iconic gesturing (as part of sign language) within the group of gestures called “sketching” (1977:67). For the gesture typology of this dissertation, the essence of sketching is condensed to the operation of Trace Leaving, \( p_{TL} \): in moving any input form through space, all the space this shape occupies at any point in time during this movement becomes part of the output form, ordinarily perceived as the ‘trace’ of this movement.

There are two types of Trace Leaving, in which the dimensions of the input form determine the output form in distinct ways, depending on the relation of the extensions of the input form and the direction of motion. When a form is moved parallel\(^{35}\) to the axis of the input form, \( p_{TL} \)

\(^{35}\) As in Mandel’s use of “parallel” (1977:67) the term is used for two forms that extend in the same direction, regardless of whether they are at a different position in space or whether they overlap.
(or parallel to the plane defined by the two axes; for 2D forms), the output form is of the same dimensions:

\[
(45) \quad p_{TL}(x_{nD}) = u_{nD}; \quad n = 1, 2.
\]

The point is excluded for parallel Trace Leaving (but see equations 50, 51), since it does not extend. A form not having an extension excludes this form from being moved parallel to the direction of its extension. Examples for one and two-dimensional input forms and parallel motion are:

\[
(46) \quad p_{TL}(\slash \slash) = \slash
\]

\[
(47) \quad p_{TL}(\rightarrow) = \rightarrow.
\]

Conforming to the operation of Continuation (equation 21), parallel motion includes curved shapes that are moved in corresponding arches:

\[
(48) \quad p_{TL}(\circlearrowright) = \circlearrowright
\]

\[
(49) \quad p_{TL}(\rightarrow) = \rightarrow.
\]

When a hand performs a gesture movement similar to (49), the hand shape has to change within the movement to remain parallel to the curved movement trajectory. In three-dimensional data, a dynamic correspondence of hand shape and trajectory possibly is an objective predictor for the operation of parallel Trace Leaving for curved output forms. All examples of Moulding (Section 8.5.3) exemplify parallel Trace Leaving.
When a form is moved in a direction non-parallel to its own axes, $p_{TLnp}$ (orthogonally or obliquely to the axes), the output form is of one additional dimension. Since a point cannot be moved parallel to its extension (it does not extend), it always falls into the category of non-parallel Trace Leaving. The simplest trace is a path (the trace of a point); a line leaves a surface as its trace; and a surface leaves a volume as its trace:

$$p_{TLnp}(x_nD) = u_{(n+1)D}; \quad n = 0, 1, 2.$$  

Expressed in graphical equations, this is

$$\text{(51)} \quad p_{TL}(\cdot_x) = \big/$$

$$\text{(52)} \quad p_{TL}(\cdot_y) = \big|$$

$$\text{(53)} \quad p_{TL}(\cdot_z) = \big\|.$$  

Since very early work on gesture (see chapters 3 and 4 in Kendon 2004), one criterion is used to categorise gestures: motion. Motion and absence of motion is objectively measurable (e.g., automatically distinguishable in motion-capture data), which advocates grouping the operations into static and dynamic gesture form operations. The static operations rest on the configuration of the articulator as for example in “independent hold” gestures, which are held still during their “expressive phase” (Kita et al. 1998:27). Still, ‘static’ does not mean that static operations would not apply to a gesture that moves during its expressive phase. It means that the operations do not need motion to take effect. They are also active in the large group of gestures that contain a stroke (see Introduction). Figure 7.11 categorises gesture form operations according to whether they are static or dynamic, giving a summary of their dimensional constraints in short-hand writing. In contrast to existing gesture categorisations (see Section 8.2), gesture form operations allow for categorisation according to specific structural aspects of form, such as the constraint on dimensionality or the involvement of motion in form creation (see also Figure 7.2 and Figure 8.59).
### Gesture form operations and Output form

<table>
<thead>
<tr>
<th>Static</th>
<th>Articulator Profiling</th>
<th>3D</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Shape Profiling</td>
<td>anyD</td>
</tr>
<tr>
<td></td>
<td>Continuation</td>
<td>sameD</td>
</tr>
<tr>
<td></td>
<td>Intersection</td>
<td>minD-1 (two input forms)</td>
</tr>
<tr>
<td></td>
<td>Adjacency (general)</td>
<td>anyD</td>
</tr>
<tr>
<td></td>
<td>Adjacency: Division</td>
<td>(n+1)D (two output forms)</td>
</tr>
<tr>
<td></td>
<td>Adjacency: Enclosure</td>
<td>(n+1)D</td>
</tr>
</tbody>
</table>

| Dynamic          | Trace Leaving         | (n+1)D      |

Figure 7.11: Overview of dimensions of input and output form across static and dynamic gesture form operations.

### 7.2.10. Alternative gesture form interpretations of a single gesture

Many gestures, often those performed in a lax fashion, are ambiguous. A single gesture can have more than one alternative interpretation of gesture form, $gf_{alt}$:

\[(54) \quad gf = gf_{alt1} \lor gf_{alt2} \lor \ldots \lor gf_{alt m}.\]

Whether one alternative interpretation is considered to be the correct one, depends not only on the specifics of bodily action (articulator form), but also on context. Some aspects of spatial context (other non-profiled articulators, other objects *etc.* ) and non-spatial context (language, memory, common ground *etc.* ) can favour conceptualising gesture form in one way; other aspects of context support a different conceptualisation. I call these aspects ‘cues’, appealing to Talmyn’s notion of a “cue” as an “element of information” conveyed in speech or gesture that helps to narrow down a target, “singling it out from alternative candidates” (Talmyn, forthcoming). According to Talmyn, the ‘target’ of an utterance may be individualised via speech and gesture, each excluding only some candidate targets by giving certain ‘cues’, while
together excluding all-but-one candidate target (Talmy 2012; forthcoming). While in Talmy’s work, gestural cues focus on targeting gestures such as pointing, here, a cue is any element of information that favours one gesture form interpretation over another gesture form interpretation of the same gesture.

In a pilot for another study (Hassemer & Winter forthcoming), I asked participants to show me the shape of the object in Figure 7.12a. Some participants conveyed the shape information by imitating enclosing the object with index and thumb (b). In a Gesture Form Analysis of this interpretation, $gf_{alt}$, the inner tactile surface of index and thumb is profiled (a ‘Holding’ gesture). This surface is adjacent to the $|object|$ to be imagined inside of this ‘O’-shape (see Figure 8.34 for a complete analysis of Holding). Yet, in discussions with some participants after the experiment, I learned that some intended to show the ring by the index finger and thumb themselves, having the full shape of both fingers roughly represent the $|ring|$ in three dimensions, $gf_{alt}$. Obviously, the size of the gesture in relation to the particular object can be a cue for which of the two interpretations is more likely in this special case, but the comparison of exact sizes is impossible with $|objects|$ whose exact shape is not known, including abstract shapes. $gf_{alt}$ and $gf_{alt}$ being different interpretations of one and the same articulator form is an example of a gesture of ambiguous gesture form.

![Figure 7.12(a,b): Stimulus object (ring) and ambiguous gesture either representing in 3D or enclosing the imagined object.](image)

7.2.11. Different articulator forms as ‘functionally homogeneous types’

Saying “I mean this book” while (a) index-finger-pointing at a book on a shelf, is an utterance unambiguous with regard to the referent. The articulators arm, hand and index finger are profiled, of which, again, the dominant one-dimensional shape is profiled, which is then continued away from the gesturer’s body to the point in space where it intersects with the
surface of the book on a shelf, identifying the book as the object referred to by the operation of Adjacency (Pointing At Object see, e.g., Figure 8.18 for a full analysis of this gesture type).

Imagine someone uttering the same sentence but (b) pointing with a flat hand palm down (fingers facing away from the body), instead of the prototypical index-finger-pointing hand. This hand configuration would have to be profiled as a planar shape, since it extends on two axes (horizontally and sagittally, on the transverse plane). The intersection of this two-dimensional form with the two-dimensional front of the bookshelf would be a horizontal line. If several books were located along this line on the shelf, gesture (b) would be ambiguous as to which book is meant, but if it was just one book (and otherwise folders) along this intersection line, gesture (b) would be unambiguous (see also Talmy 2012:11-2; forthcoming). Pointing with the index finger (a) and pointing with the flat hand (b) would single out the same book—although with different articulator form and by different gesture form interpretations: The line profile affords pointing to the same book unambiguously as does the plane profile in a certain context.

This shows that two alternative combinations of constituent forms can afford the same gestural function of singling out one specific book by pointing—in spite of the constituent forms being different. For a setting with just one book on the shelf, gestures as (a) and (b) are called ‘functionally homogeneous types’.

7.2.12. Conceptually reducing the dimensions of gesture space

For both the gesturer and the interlocutor facing the gesturer, identifying sizes and shapes that extend sagittally is more difficult than identifying sizes and shapes that extend vertically or horizontally in the frontal plane. Displaying measurements sagittally is thus often marked and used only when specifically needed, for example when three consecutive gestures display height, width, and also, on the sagittal axis, the depth of an object. Generally, the space in front of the body does seem, by default, to be treated as three-dimensional. For example, when gesturing as if to hold something between the opposed vertical flat hands held in front of the body, the space between the hands can be conceptualised as three-dimensional (equation 42; Figure 8.30). Still, this does not need to be the case. Whenever the sagittal axis does not seem to be of importance, gesture form can also be imagined in two-dimensional space. The same gesture of opposed vertical flat hands held in front of the body may be interpreted as

---

36 For more information on the different roles of index-finger and flat-hand pointing in discourse, see Kendon 2004 or Fricke 2007.
delimiting only a two-dimensional area in a frontal plane. In two-dimensional gesture space, the hands’ surfaces would be reduced to vertical lines that enclose an area of the plane (equation 39). The next possible reduction of gesture space is to one dimension, as often occurs with exactly the same gesture to indicate the size of an object on one axis only. In this gesture space, the hands’ surfaces would be reduced to two points that enclose a line segment, the object’s width (equation 38).

Reducing the dimensions of gesture space should not be confounded with merely reducing the dimensions of a single constituent form. The dimensions of gesture space can be reduced per se. As introduced in this chapter, the dimensions of gesture space are Euclidean. Thus, a low-dimensional gesture space projects all forms of the three-dimensional gesture space onto its (remaining) axes, reducing gesture space to a two-dimensional plane or a one-dimensional line. This has an impact on all constituent forms, potentially reducing all their dimensions by projecting orthogonally onto the new, low-dimensional space. In a two-dimensional space, only forms of zero, one or two dimensions are possible. The reduction impacts not only the dimensions of form but also the affordances of forms because space itself is reduced. This means that forms of lower dimensions can divide and enclose space as laid out in Figure 7.4 and Figure 7.5. By reducing the dimensions of gesture space, the entire spatial arrangement gets reduced, often simplified. Gesture space reduction onto the transverse plane or a horizontal or vertical line will be empirically exemplified in Figure 8.42, Figure 8.54, and the accompanying analyses.

7.2.13. Where gesture form ‘begins’

This section treats the problem of determining—within the progression of cognitive activity of gesture perception—at which point a Gesture Form Analysis should begin, and how. In this dissertation, the first two gesture form operations, Articulator and Shape Profiling, are treated as two distinct operations, but they may also be conflated into a single operation. There are different perspectives on how they influence gesture conceptualisation, touched upon in the following.

The two operations Articulator Profiling and Shape Profiling are grouped under the label of ‘profiling’ operations, for a particular reason. Their effect on gesture form could just as well be combined to one single operation, since they always occur in the same succession and they have complementary functions of profiling first a part, in the sense of a three-dimensional
portion, of the body (Articulator Profiling) and then profiling yet another part, in the sense of salient shape aspect, of this form (Shape Profiling).

With regard to the operation of Articulator Profiling, the present theoretical approach cannot make assumptions on whether an observer actually first perceives the whole body with all its potential articulators and only afterwards profiles a certain articulator unit. Among all the potentially meaningful parts of the visually accessible scene, attention is focused on certain shapes, often the articulators of the interlocutor, because they could convey information—without regarding all potentially relevant shapes. In consequence, Articulator Profiling might not be an operation that selects one among many potential articulators, but an operation that identifies one visually salient articulator unit to convey consistent information so that the form information of other articulators is not even considered. This interpretation is a viable alternative since it could mean that Articulator Profiling is a way to reduce processing load in visual perception.

With regard to the operation of Shape Profiling, this approach cannot make assumptions about whether an articulator is ever perceived in its complete three-dimensional complexity. If the flat hand was simply perceived as two-dimensional (or a straight articulator as one-dimensional) in the first place, no Shape Profiling would be needed. Whether Shape Profiling reduces the perceived three-dimensional shape to a shape of fewer dimensions or whether Shape Profiling actually means limiting the perception of an articulator to its predominant extent, is beyond the scope of this dissertation. Again, the operation of Shape Profiling could play out to be a way of omitting processing load in visual perception and conceptualisation. Regardless of whether the operations Articulator and Shape Profiling base on a full perception of a three-dimensional body or whether they are a strategy to reduce processing load, they produce one and the same output form, consistent with the proposition of Gesture Form Analysis.

Both profiling operations mark the beginning of gesture form interpretation. They account for how a specific part of the body is profiled, from which the spatial information of a gesture is derived. Articulator and Shape Profiling are to be understood as abstractions that further process the specific aspects of the rich information of the gesturer’s whole, dynamic three-dimensional body. Consonant with theories of perception, these operations could profit from the perceptual process that focusses on lower-dimensional form aspects as they are part of the perception of three-dimensional forms: since the objects’ two-dimensional surfaces are an intermediate step of visual perception (Marr’s 2 ½ D sketch, see Section 4.2.2), reducing the
three-dimensional form of the articulator again to these surfaces could require only little
cognitive load (since the objects had just been represented by these surfaces). The operation of
Shape Profiling might also be entrenched with later processes of visual processing, the object-
centred representation since, for example, in pointing gestures the direction of the articulator
is crucial—not from a viewer-centred perspective but from an object-centred perspective (in
this case articulator-centred) — to find out what the gesture points at. This includes, for
example, correcting “the foreshortening of an object’s principal axis of elongation” (Ellis &
Young 1996:39; see also Marr 2010/1982:313–8; a recognition process impaired specifically in
subjects with right-hemispheric lesions, Warrington & Taylor 1973).

In conclusion, (1) Articulator Profiling and Shape Profiling are conflatable. (2) It cannot be
determined what apperceptive or associative processes of visual perception are reflected
specifically by these profiling operations, but there seem to be several processes, spanning
from early to later stages of visual perception, that are consonant with the operations
identified as necessary in Gesture Form Analysis.

7.2.14. Where gesture form ‘ends’: the resulting form

As repeatedly mentioned, how an observer conceptualises gestures is highly dependent on
context. In order to focus on the structural details of a theory of gesture form, this context has
not received only theoretical (Section 7.2.10), but no empirical attention up to this point. The
following chapter will therefore go through a spectrum of examples, situating gesture form in
different kinds of context. In preparation for this empirical chapter, this section considers the
structural question of whether there is a most important, ‘last’, ‘final’ or ‘resulting form’ among
the series of constituent forms analysed by the approach at hand.

Which constituent form is perceived as the resulting form? The decision about which
constituent form is perceived as the part of spatial information whose conveyance is central to
a gesture (the resulting form), depends on the setting of the communicative act. For example,
the context of object descriptions biases an observer to focus specifically on object features
that could be conveyed by the gestural performance. Thus, the study conducted for this
dissertation (Study 2, Part III) is inadequate specifically for supporting claims about
mechanisms that determine the resulting form. Yet, some gestures are less context dependent
with regard to the resulting form than others. A theoretical example of such a gesture is a
pointing gesture, in which the resulting form is a point (the indicated location): though other
forms are necessary to understand this gesture (as is the core claim of this dissertation),
observers of this gesture will be able to report on the location pointed at. This is an example of a clear end of Gesture Form Analysis and a clear resulting form. Another theoretical example is a person moving an index-finger-pointing hand from left to right in a slow, controlled movement. Observers will likely identify a line being drawn in the air. The form of a line then is the resulting form of this gesture. Not all observers will readily name a specific form as most important in a given gesture. In the first example, they might just say “oh, she pointed over there”. Nevertheless, the information that they attribute to a given gesture is based on spatial information about the resulting form: the last constituent of gesture form. Gesture types, as for example Pointing At Location, consist of several constituent forms (e.g., Figure 8.14). Still, observers mostly identify only one resulting form: the last form in the Gesture Form Analysis. As this seems to be the standard case (pending experimental quantification and support), the last of the constituent forms is to be understood resulting form, unless indicated otherwise.

In some gestures, people differ in what they regard as the resulting form. When someone mimes opening a window, (1) some observers may say that this gesture shows how one opens a window, focussing on the movement of the hand. (2) Others may interpret that it is to show how a certain window opens; focussing on the |window| and the way it rotates or slides when opened. There are thus two different resulting forms that may get marked as the core information within one and the same sequence of constituent forms. The first group of observers perceives the profiled hand and its movement (articulator form) as the resulting form of the gesture, the second perceives the |window| that is adjacent to the surface of the profiled articulator unit to be the resulting form. While both conceptualisations will most likely include imagining the same constituent forms, that is, how the hand manipulates the window that is in surface contact with it, interpretation (1) already sees the articulator form to express the gesture’s core spatial information. It identifies the movement of the hand during the expressive phase as the resulting form. Interpretation (2) foregrounds the use of the operations Shape Profiling and Adjacency inferring the |window| that is adjacent to the surface of the articulator as the resulting form. Note that interpretation (1), in an extreme case, does not necessarily include conceptually processing the gesture beyond the articulator form (for example an interior designer using this gesture to illustrate that the space in front of the window should be left empty because the hand and arm need space to open it). For such an observer, gesture form could ‘end’ at the manual movement, while for the observer in (2), the gesture necessarily includes imagining the form of the window—at least where in contact with the hand.
This section is to highlight the fact that different observers interpret different constituent forms to be the core spatial information. Even one and the same observer could be undecided between more than one possible resulting form. The practical analyses will not focus on different resulting forms but will rather analyse all constituent forms regardless of their relative importance, focussing on the overall organisation. Speech and other non-spatial context, as well as spatial context, certainly contain cues for the prominence among the constituent forms within one gesture, but this is beyond the scope and beyond the methodology of this dissertation.

7.3. A definition of gesture form

Gesture form is an observer's conceptualisation of the spatial information conveyed by a gesture. Gesture form is defined as the output of a combination of gesture form operations on the gesturing body and consists of forms of zero to three dimensions. There is a total of eight gesture form operations, of which the first two occur in the beginning and in a constant order. The remaining operations are optional and can occur in any order.

Of the gesturer's body, certain articulators are profiled of which, again, forms of three, two, one or zero dimensions are profiled (1: Articulator Profiling, 2: Shape Profiling). Other static (3: Continuation, 4: Intersection, 5: Adjacency, 6: Division, 7: Enclosure) and dynamic (8: Trace Leaving) operations add other constituent forms. Thus, each gesture contains multiple constituent forms, beginning with the gesturer's three-dimensional body. The operations then elicit constituent forms of the same number of dimensions (operations 1, 3), of one additional dimension (6, 7, 8) or of a differently calculated number of dimensions (2, 4, 5).

The notion of gesture form structures the elements of spatial information conveyed in a gesture. It is a basis of gestural meaning. Gesture form can stand for itself (the direction in a route description) or be mapped onto spatial reality (showing the size of a ball) or abstract concepts (moulding a round shape to describe a well composed 'round' story). Spatial and non-spatial context, including speech, influences the conceptualisation of gesture form.

7.4. Gesture form operations as a spatialisation of semiotic relations

There are coarse correspondences between gesture form operations and semiotic relations used in gesture literature (i.a., Cienki & Müller (Eds.) 2008; McNeill 1992; Mittelberg 2010b). Figure 7.13 will sketch correspondences between the two. The ensuing chapter will exemplify the use of gesture form operations instead of semiotic notions in qualitative analyses. This
dissertation proposes specific, that is formalisable, spatial operations, which structure the way gestural action is conceptualised. This is possible because gesture uses the various affordances of space ‘iconically’ (to use established terminology)—perhaps more iconically than spoken and written language. I propose that a semiotic notion (describing a relation between a semiotic object and a semiotic sign), such as iconicity, can be taken to imply multiple spatial relations, more precisely gesture form operations, which describe the relation between an input form and an output form as part of gesture conceptualisation.

The goal of this dissertation, then, is to investigate the spatial implications of iconicity in gesture (or ‘iconic devices in gesture’, following Mandel 1977) as well as other semiotic notions. It is taken as a given that gestures can be iconic with regard to their referent, yet Streeck’s (2009) criticism highlights a research gap (that this dissertation desires to contribute to): iconicity is too broad a term to specify the relation of the hand and the reference object. In fact, almost all types analysed in this dissertation fall in the category of ‘iconics’ in that the depicted forms share form aspects with the articulating body part, or derivations thereof. Gesture Form Analysis asks the question “In what respect is the articulator iconic?” or “In which way does the articulator overlap with the depicted form?” Answers to these questions vary considerably. A gesture can depict the height of an object (similarity with respect to the extent on one axis only, Figure 8.32), it can show the motion of an object (similar with respect to the motion trajectory, Figure 8.21a), it can indicate the location of a referent (the location indicated is similar or equivalent to the referent’s, Figure 8.16), and so on. Iconicity is, with regard to gesture form, an umbrella term that includes many of the ways in which the conceptualisation of gestural action infers a referent. This referent can be a physical object, as in many examples taken from the dissertation studies, or an abstract concept, as in many of the cited and analysed gesture examples taken from gesture literature (see Chapter 8).

Gesture form operations can be seen as geometric specifications of semiotic relations. Thus, they are not competing with semiotic relations. Rather, they are based on semiotic relations. Gesture form operations attempt to describe how semiotic relations manifest as spatial relations in the conceptualisation of a gesture. The spatialisation of certain semiotic notions is rather straight-forward or has already been described in gesture literature. The notion of external metonymy has clear spatial implications. It corresponds to operations that are tightly related among each other (both Division and Enclosure are specifications of Adjacency, see Sections 7.2.6-8). For example in McNeill’s cartoon gesture’ (Figure 2.7): “the imagined objects are adjacent to (contained in or sitting on), but external to, the hands” (Mittelberg & Waugh 2009:340). Since the hands in this example enclose the imaginary object
on both sides, the operation of Enclosure as part of a Gesture Form Analysis can almost fully rely on the concept of external metonymy, merely specifying the form dimensions of the involved forms (Figure 8.30). Other semiotic relations do not have a clear match in a gesture form operation but apply to multiple divers aspects of gesture form creation at different stages of conceptualisation, for example iconicity, indexicality or internal metonymy.

Figure 7.13: Gesture form operations and approximatively corresponding semiotic notions.

Figure 7.13 can only be approximate in attributing semiotic notions to gesture form operations; it includes semiotic relations of different hierarchical ranks and clearly cannot claim exhaustiveness. The figure illustrates that notions such as indexicality can be broken down into more than one spatial relation and therefore also into more than one gesture form operation: Continuation, Intersection, and Adjacency (in its three variants). Empirical examples (Chapter 8) will show these operations to occur in different gesture types and independently of one another. This is, again, consonant with semiotic relations, which also do not exclude each other within one and the same semiotic sign, but co-occur in a single word or a single gesture (Ransdell 1986; see also Section 2.6). Figure 7.13 further shows kindred semiotic relations (e.g., iconicity and internal metonymy) relating to a similar range of operations.

Gesture form operations are claimed to describe all aspects of gesture form—as far as this can be said in the light of the limited number of examples in this dissertation. Follow-up studies have to test the typology’s ability to capture additional empirical gesture data. The operations only approximately corresponding to semiotic relations emphasises the necessity to introduce a new theoretical approach including new terms. If, for instance, Continuation was
just another word for indexicality, there would be no need for a new terminology. But the
operation of Continuation is *more general* than indexicality in being applicable to more
complex forms than only straight lines. It also applies to curved lines, as well as (curved)
surfaces. Continuation applies to a range of different forms, thus being part of different
gestural functions. At the same time, Continuation is *more specific* than the linguistic notion of
indexicality: there are gestures, for example pointing at a certain person (Pointing At Object
(no contact), see Section 8.4.6), which combine Continuation and Intersection and Adjacency.
All of these operations are potential derivatives of the notion of indexicality, but there is no
fixed link between Continuation and Intersection and Adjacency: Other gesture types are
structured merely according to the operations of Continuation and Intersection (Pointing At
Location, Section 8.4.3), or specifically by Continuation (Directing, Section 8.4.2). Yet other
gestures combine operations related with indexicality with one or more operations associated
with other semiotic relations (e.g., any of the gesture examples given so far in this dissertation
include iconicity in the sense of the articulator exhibiting similarity to a simple geometric
form). The combination of different ways of Profiling, Intersection and Trace Leaving can
make Continuation a building block of a very different gesture: Drawing (Section 8.4.7).

In addition to the vague correspondence to semiotic notions, the proposed gesture form
operations may be associated more precisely to spatial relations such as spatial part-whole
relationships (part of the broader notion of internal metonymy), different kinds of spatial
adjacency (see also Sections 7.2.6-8), and spatial metaphor, as illustrated in Figure 7.14. The
operation of Trace Leaving is here described as a metaphor, since for example drawing a line in
the air actually does not immediately present the line to the eye of the observer. Instead, only
the interpretation of the performed motion as standing for the spatial extension of a (static)
line, allows perceiving this line as the spatial information conveyed in this gesture (see also
Mandel 1977 and Section 2.5). Motion through space and spatial extension are “different kinds
of things” (Lakoff & Johnson 1980:5), motion being the source and spatial extension being the
target domain, which makes this mapping a metaphor. Since both source and target of this
metaphor are spatial, I would like to call it a ‘spatial metaphor’.
Chapter 8 will contain the empirical analyses of gestures by combinations of the gesture form operations introduced in this chapter. This will break down gesture types and define them more precisely with regard to form. For example, the gesture type 'Pointing At Object (no contact)' stands for the combination of a one-dimensional Shape Profile, Continuation, Intersection and Adjacency. These four structural constituents are constant among pointing gestures of the type Pointing At Object (no contact), but also recur (in different combinations and order) as building blocks of other gestures types (not only pointing gestures). The introduced graphical coding scheme (Section 7.2.2) will be employed in each of the following analyses. The goal of this coding scheme is to present the gesture form operations (and the constituent forms they create) as a structure for visualising a gesture's necessary processing of spatial information.
8. Specific Gesture Form Analyses building up to a gesture typology

As starting point for the empirical part of this dissertation, Section 8.1 presents an exploratory study of gesture form with Study 1, which served as a pilot for the empirical testing in Study 2 (Study 2 is introduced at the beginning of Part III). Study 1 includes gesture data of participants completing tasks that deal with diverse kinds of spatial information. The data contains diverse gestural types that will add to the gesture typology at the end of this chapter. The beginning of this chapter will touch upon existing functional categorisations at the example of two related gesture types, for whose differentiation the proposed structure of gesture form may contribute (Section 8.2). Section 8.3 will discuss the approach of analysing diverse, selected gestures in its potential and its limitations. The following three sections are arranged according to the main classes of the gesture typology, each beginning with structurally simple gestures and build up to gestures of a more complex structure. Section 8.4 analyses gestures, in which the articulator unit is profiled as a line. In Section 8.5 the articulators are profiled as a surface and in 8.6 as a volume. To slowly unfold the complexity of the gesture typology, each of these three sections includes gestures of similar functions. Complementarily, Section 8.7 deals with the remaining gesture types, in which form profile and function do not necessarily align as in Sections 8.4-6. This section shows the flexibility of the matrix: form profile and other constituent forms combine in manifold ways, affording a variety of gestural functions. The typology in Section 8.8 summarises recurrent gestures in a matrix of primitives (dimensions) and gesture form operations.

8.1. Study 1

Particulars of the methodology of Study 2 will be postponed to Part III of this dissertation, only Study 1 will be introduced in this part. This is because Study 2 is designed to collect two specific gestures for experimental hypotheses (Chapter 9). Study 1, in contrast, spans a variety of tasks, eliciting many different kinds of gestures. This broad range of gestural data (in addition to gestures discussed in gesture literature) is to be used for the purpose of proposing a typology for all gestures, as endeavoured in Chapter 8.

In the first session of Study 1, one participant is recorded while describing twelve objects (Figure 8.1). In a second session, another four participants describe the objects themselves, and are also asked to report what they see in the approximately 20-second long, mute video clips.
recorded from the first participant. They then guess which object of the ones laid out in front of them is referred to. Further, the participants are to fulfil multiple other small tasks. Among these tasks are route descriptions of their way to the Natural Media Lab, both through the city of Aachen and through the complicated system of elevators and staircases within the building. Other tasks include estimating cardinal directions and positions of different parts of the city and counting the numerous cameras and reflective markers (that are used for optical 3D motion tracking) in the room.

In its diverse tasks, Study 1 covers different domains of spatial reality that gestures are studied in, thus providing a broad range of examples to illustrate the functioning of operations introduced in Chapter 7. The evaluation of the object description task informed the design of Study 2, which aims at testing specific hypotheses with motion-capture data (Chapter 9).

8.2. Some differences to functional gesture categorisations

Before going into analyses of gestures that exemplify the different types of a strictly form-based gesture typology, this section touches upon some pertinent categorisations that exclusively or partly use functional criteria for their distinctions. A comprehensive review of categorisations (including also, e.g., Efron 1941; Fricke 2007; Sowa 2006; Streeck 2008) would depart too much from the focus adapted in this dissertation. Here, the focus is on a different kind of analysis, whose modular nature outputs a typology by listing the required spatial operations for each type (Figure 8.59). This is not to say that functional categories and the form-based distinctions proposed here could not complement each other, but merely that the aim of this part of the dissertation is specifically to introduce a typology based on distinctions that can be broken down to a fix set of form differences. This is a core distinction between functional categorisations and the proposed typology. Not only the proposed categories differ, but more fundamentally, the criteria for distinction are not the same. Thus, a comparison of all
categories will not be carried out, but this section will be limited to exemplifying a characteristic structural difference.

The brief overview in Figure 8.2 illustrates a very divergent treatment of two types of gestures, pointing and drawing, in several gesture categorisations. Though inevitably simplifying, this figure captures a structural feature of these categorisations. Pointing (or “indexical presentation”, “deictics”) and drawing (or “tracing”, or part of “sketching” or “iconics”) are grouped with very different gestures. Pointing, for Kendon and McNeill, stands on its own, for Mandel it shares a category with “mime” and for Enfield with “placing”. Drawing groups with “modeling” and “enactment” (Kendon), “stamping” (Mandel) or remains on its own (Enfield, McNeill, Müller). The grouping and the hierarchical organisation do not show agreement on which other gestures the two types are related to. On this basic structural level, agreement can be found only in the fact that the two types never occur in the same sub-category—on the contrary, they are always set apart by the most basic hierarchical distinction.

This contrasts to the close relationship between pointing and drawing (with regard to gesture form), as described for example by Goodwin (2003a:229-230): “Pointing gestures can trace the shape of what is being pointed at, and thus superimpose an iconic display on a deictic point within the performance of a single gesture.” As discussed in Chapter 3, the similarity between pointing and drawing might be captured by the imaginary forms that are part of their conceptualisation. Thus, the intended contribution of this chapter is to go through the forms that are necessary for the conceptualisation of instances of one gesture type and to identify the similarities and deviations of related types, such as in the relation between pointing and drawing. The result of a so-structured analysis can be a tree structure, which illustrates the overlap in gesture form, while also isolating the constituents that set drawing gestures apart from pointing gestures (Figure 8.20), or it can be a table, emphasising the modular nature of the typology (Figure 8.59).

37 Picking out pointing and drawing in these categorisations is motivated by having already investigated their similarity in forms within gesture analysis of various authors (Sections 3.2 and 3.3).
8.3. Developing a typology by systematic gesture analyses

In functional categorisations, the communicative goals determine to which category a gesture belongs. In contrast to this focus on functionality, I would like to introduce form affordances underlying Gesture Form Analysis. On a fine-grained level, form affordances are form-function mappings. Affordances are, on this level, a direct result of form (see Section 7.1.2). Gesture Form Analysis assumes that if observers cannot understand the form that is to be abstracted from bodily action (gesture form), they cannot understand its communicative function. Different kinds of form (affording the communication of different spatial information) are thus the most basic distinction. Figure 8.3 illustrates this interdependence: different kinds of form afford structuring space in a meaningful way. The so structured space contains different forms (output forms) that have a specific relation to the kind of form that imposes the structure
(input forms, "kind of form 1" etc.). The relation between the input form and the output form and can be specified (and even formalised) by certain geometric operations. I would like to use these geometric operations as the building blocks of the conceptualisation of space in gesture interpretation.

![Diagram](image)

Figure 8.3: Diagram showing spatial operations (3rd level of the hierarchy) as a result of specific forms (1st level) affording specific spatial structures (2nd level).

One operation does not suffice to capture the spatial information that gestures convey. Rather different operations are combined, and the combination of operations defines gesture types that can be described by a typology as illustrated at the end of this chapter (Figure 8.59). This figure shows the theoretical proposal of this dissertation in a nutshell. It displays the typology in the form of a matrix whose modular design is one way to explain the incongruous categorisation of pointing and drawing gestures in the above functional categorisations.

A structure defined by the two principles of form dimensions and gesture form operations as laid out in Chapter 7 theoretically allows for an infinite number of combinations of form operations corresponding to an infinite number of possible gesture types. A typology of gesture cannot (and should not) address all theoretically possible combinations of operations. The gesture types in the typology presented here include all combinations of operations that were necessary to describe the gestures found in Studies 1 and 2 and in the cited literature. Surely, gestures of a different context might require new combinations of operations that I do not address in my typology. The aim of Gesture Form Analysis is to provide structural elements that, in new combinations, can account for gesture types that this dissertation does not cover.

As a terminological clarification before going into practical examples, note that there are distinct terms for (1) ‘gesture form’ versus (2) the ‘structure of the gesture form’. (1) One ‘gesture’ includes various specific ‘constituent forms’, which have specific ‘form qualities’. Each specific ‘gesture’ is a token of (2) a ‘gesture type’, also called a ‘combination of operations’ or a
‘structure of gesture form’. The following four sections will provide a range of Gesture Form Analyses that show the structure of gesture form, as well as the size, position and motion of their constituent forms in a graphical illustration. To emphasise structural commonalities between types, sections comprising several related types are grouped together, beginning with the least complex and ending with the most complex (e.g., Section 8.4.1-8.4.7).

8.4. Pointing and Drawing

Gestures understood as pointing gestures are regarded as indicating an object or a location that is discovered by projecting a straight line from the furthest point of the body part that has been extended outward into the environment. We shall not here enter into a discussion of this process of interpretation, how it arises, or how individuals come to adopt it. To do so would take us too far from our present aims. We would like to point out, however, that how the object or location that intersects with this straight line is identified as being the object or location referred to by the gesture is not well understood. It clearly depends on what else is being said or done as part of the utterance of which the gesture is a part, and also on the presuppositions shared by the interactants, in terms of which what is relevant and what is not relevant for the discourse are understood.

Kendon & Versante 2003:112

On one hand, the perspective taken in this quote, including part of its terminology, overlaps considerably with the proposed Gesture Form Analysis. On the other hand, my investigation is also complementary to Kendon and Versante’s in that I will specifically look into the process of interpretation that includes a straight line extending outward from the furthest part of the body and the intersection at a certain location or object. These and related aspects of a form interpretation will be investigated by means of examples from the gesture literature and from Studies 1 and 2. The gestures taken as examples span different functional classes, for example deixtics, iconics, and metaphorics (McNeill 1992). This is to propose that all these gestures, whose form-function mappings are not entirely arbitrary and conventionalised (see Section 2.6), are analysable with regard to gesture form.\(^{38}\)

Chapter 3 reviewed analyses in the gesture literature and found specific forms being used to describe for example pointing and drawing gestures. The forms of line and point suggested a

\(^{38}\) Also included are so-called beats and cohesives (McNeill 1992), but their gesture form is drastically abstracted in shape, sometimes consisting of movement or even ‘indefinable activity’ only (Sections 8.7.6-7).
structural relation of pointing gestures as reiterated in Figure 8.4. This chapter will underpin this rough structure by breaking several pointing types down to specific operations that decide over the dimensionality of all the forms constituting these gestures.

<table>
<thead>
<tr>
<th>&quot;line&quot; extending from articulator</th>
<th>plus: &quot;point&quot; at intersection</th>
<th>plus: moving, the point &quot;leaves a trace&quot;</th>
</tr>
</thead>
<tbody>
<tr>
<td>&gt; directing</td>
<td>&gt; pointing at a location</td>
<td>&gt; drawing</td>
</tr>
</tbody>
</table>

Figure 8.4: Simplistic hierarchical distinction of gestures that direct, point and draw.

To illustrate a structure that entails the proposed operations (and the gesture types they add up to) as well as a relational hierarchy of types, every section will be headed by a relational chart that shows the operations as columns and types as branches in an expanded tree-diagram showing their hierarchical relation as in Figure 8.5. Each gesture type contains all the operations of the types to its left. The use of upper case initials as in ‘Pointing At Location’ indicates a type that is precisely defined in its gesture form operations. Although several variants may exist, these variations are also precisely defined in terms of the constituent operations. Figure 8.20 exemplifies an elaboration of the combination of a tree-diagram and columns as introduced in Figure 8.4.

8.4.1. The basis for Pointing and Drawing: Representing in one dimension

Leading to the group of pointing gestures, there is a gesture type that only consists of profiling the articulator as a line. This type will be shown to contain the conceptual operation that most pointing gestures build upon.

---

39 The operation of Articulator Profiling is not a distinctive part of the analysis since it occurs in all gestures analysed in this dissertation. The distinctive operations being independent as actually an
In an example by Enfield, a Lao man “is reporting on a problem in the installation of drainage pipes from a bathroom block [...] ‘Make it steep like this’” (Enfield 2009:6). While uttering the quoted sentence in Lao, he extends his right arm and hand with the palm rotated upwards as shown in Figure 8.6a, while looking at them. According to Enfield (ibid.), the man’s arm and hand stand for the recommended steep angle of the pipe. Since this gesture specifically deals with this steep angle recommended for the construction, the shape and orientation of the man’s arm and hand are relevant in their angle only—or they are at least strongly foregrounded.

To understand this gesture as an indication of a specific angle, an observer has to profile the gesturer’s right arm and hand (“Articulator Profiling” in Figure 8.6b). This articulator unit has to be understood not in its three-dimensional complexity, but rather as a simple, one-dimensional shape carrying the information of angle to be conveyed as gestural contribution to the overall utterance (“Shape Profiling” in Figure 8.6b; see Figure 8.14a for another example of Representing in 1D, in which a finger stands for a street).

---

important claim of Gesture Form Analysis (see Section 7.2 or “multiarticulatory gestures” in Hassemer 2009).
If, alternatively, the articular unit hand and arm was to stand for the drainage pipe also in the round circumference of the arm (exemplified in Enfield 2009:19), or in the plane shape of the flat hand palm up, a different interpretation of gesture form would be in place (namely Representing in 3D, Section 8.6).

Figure 8.6(a,b) and all the following figures of the same kind are the graphical representation of Gesture Form Analyses. They comprise an original picture and a zoomed-in picture onto which the colour-coded gesture forms of the respective example are superimposed. Attached right underneath, the coding scheme (introduced in Section 7.2.2) and its according operations are to be found. They are colour-coded, too. The overlay of coloured forms in the picture has the same colours as the respective shape descriptions (including their dimensions) in the coding grid. The three-dimensional forms of the body and the profiled articulator unit are always black in the coding grid; other abstract, zero to three-dimensional forms are colour-coded according to the operation that they emerge from. In Figure 8.6, only the one-dimensional profile of the arm and hand is marked by a green line. Figure 8.7 displays all colour-coded gesture form operations, whose colour will be maintained regardless of the form dimensions and the order of operations (see e.g., Figure 8.24).

**Shape Profiling, Continuation, Intersection, Trace Leaving**

Figure 8.7: Four gesture form operations, which are displayed with the same colours throughout all graphical Gesture Form Analyses of this dissertation.

8.4.2. Directing

Figure 8.8: Excerpt of a typology of pointing gestures in the form of a tree-diagram, including Directing as an immediate elaboration of Representing in 1D. (The following figures at the beginning of each section will build up on this diagram. The bold-faced dimensions below the name of the operation show the number of dimensions of the output form determined by that operation and the preceding ones. The operation of Articulator Profiling is not part of these diagrams since it occurs in all gestures analysed in this dissertation.)
This section introduces the next gesture type that can be seen as an elaboration of the last type. In an example from Kita (2003b:308-319), a pedestrian in front of the library of a university in Tokyo instructs her interlocutor to go diagonally-left as part of a route description to a book store in walking distance. When she utters “left”, she performs a gesture with her left hand that indicates a direction (Figure 8.9a). This gesture will serve as an example for ‘Directing’, and it requires three gesture form operations (Figure 8.9b), one operation in addition to those in the last example (Figure 8.6b). First, certain articulators are profiled: the stretched-out arm hand and index finger of the left hand. They form an articulator unit that gets profiled as a straight line segment (green line in Figure 8.9b), since they predominantly extend along one (straight) axis. This segment is continued in a straight line that extends infinitely in one direction (orange line). This line is the core of the spatial information conveyed by Directing. The interlocutor not looking onto the gesturer’s articulators themselves but rather in the pointing direction where the continued line, the here-proposed resulting form of the gesture, is located, supports the analysis, especially the existence of the imaginary continuation of the profiled line.

The mentioned three geometrical operations (Articulator Profiling, Shape Profiling, Continuation) are a necessary condition (see beginning of Chapter 7) for understanding this gesture as Directing. The lack of just one of these operations—and this argument is what this chapter is to show for all analysed gestures—makes it impossible to understand a gesture of this type. If either (1) the index finger, hand and arm were not recognised as the profiled articulator unit in this gesture, or (2) this articulator unit was not identified to predominantly extend along one axis, or (3) it was not understood, that the line segment profiled from the extended articulators (until now, the spatial information conveyed in this gesture is an angle only, as in Figure 8.6) had to be continued away from the gesturer so as to function as indicating a direction in which the interlocutor is supposed to walk, the body movement would not afford conveying a direction as it is the case for this part of the route description. Each of the three spatial operations is a necessary constituent of conceptualising Directing.
The reason why the participant raises her arm (instead of holding the articulator unit parallel to the ground) may be to ensure that the gesture is visible to the interlocutor, especially since the interlocutor is not looking at her but in the direction indicated, as is the case at this moment. The vertical inclination of the hand and arm does not seem to be part of gesture form; it does not seem to be relevant in conveying the direction (in contrast to Figure 8.6). Rather, the space in which the gesturer conveys directions to the interlocutor can be reduced to two dimensions, a horizontal plane like on a map, in which height information is backgrounded (see also Fricke 2007:266–7 for ‘map-like’ ("kartenähnliche") reductions of space). Analyses of Figure 8.32 and Figure 8.42 treat the reduction of gesture space dimensions systematically.

Figure 8.10a shows another gesture of the type Directing in a different setting, for a different purpose and performed with a more reduced articulator unit. As a response to the question “Where is North?” in Study 1, the participant’s gesture of a raised index finger pointing in a specific direction has to be conceptualised by same combination of operations (Figure 8.10b). Again, the vertical inclination of the articulator unit seems to be irrelevant.

---

40 This is not to exclude the possibility that distance might be coded in the angle of the articulator unit as argued for Arrernte speakers in Wilkins 2003:188–9.

41 The participant’s arm is not part of the profiled articulator unit that indicates a direction, but only the index finger (the difference in angle between arm and hand/finger is bigger than it may appear from this camera perspective).
Only the horizontal inclination is relevant for interpreting her response to the task of indicating a cardinal direction.

Figure 8.10(a,b): Directing a cardinal direction in response to the prompt to indicate where North is (@nmLab).

Among the different types that fall into the group of pointing, Directing contains the simplest combination of operations, that is, employs the smallest number of gesture form operations (Figure 8.20). The following pointing types will contain not only the operations that were argued to be a necessary condition for understanding the gesture type Directing (Articulator Profiling, 1D Shape Profile, Continuation), but also additional operations.

8.4.3. Pointing At Location (contact)

Figure 8.11: Excerpt of a typology of pointing gestures, including Pointing At Location (contact) as a elaboration of Representing in 1D (an alternative elaboration to ‘Directing’ in Figure 8.8).

Starting with an articulator conceptualised as a line, another operation, Intersection, adds to a different gesture type, Pointing At Location, as in an example taken from Andrén’s (2010) analysis of children’s gestures. As illustrated in Figure 8.12b, Continuation is not necessary,
since the articulator is in contact with the referent (the operation of Continuation is ‘skipped’ in Figure 8.11). For Pointing At Location (contact), gesturers have to be touching the location they are pointing at.

As Figure 8.12 illustrates, the line that the articulator represents, meets the surface of the table, on which parts of a puzzle are placed. With the help of this gesture type, the boy indicates the location where to put the puzzle piece in question. The verbal contribution “there” (“där”) does not convey spatial information but rather emphasises that the gesture is doing so. Note that the location indicated is unnecessarily precise. Another point within the area the puzzle piece is supposed to cover would also be acceptable. Still, a central point might be best for unambiguous interpretation.

Figure 8.12(a,b): An infant Pointing At Location (contact). Adapted from Andrén (2010:86; excerpt). Courtesy of Mats Andrén.

Even if a larger articulator unit is profiled here (index finger, hand, arm), which would have to be interpreted as a curved or bent line, the same point in space results from the intersection of this line and the surface, since articulator and surface are in contact. This also speaks for the ease of pointing with contact. In the type analysed in the next section, a curved shape profile would alter the location pointed at.

Mandel calls the gesture type in Figure 8.12b “stamping”, in which “the implement moves forward and then returns, like a rubber stamp” (Mandel 1977:67). Like in Gesture Form Analysis, his approach is open to have not only a line but also other forms, for example another surface, intersect with the surface of the objects stamped on.
8.4.4. **Pointing At Location (no contact)**

![Diagram of pointing gestures]

Figure 8.13: Excerpt of a typology of pointing gestures, including Pointing At Location (no contact), which combines the operations of Continuation and Intersection.

Apart from profiling an articulator as a line, Pointing At Location (no contact) involves combining Continuation (the distinctive operation of Directing), and Intersection (the distinctive operation of Pointing At Location (contact)), as illustrated in Figure 8.13. The discussed types of pointing gestures build up to the complexity of the most prototypical way of pointing, Pointing At Object\(^{42}\), which will be described in the next two sections.

Figure 8.14a shows a gesturer giving a route description on an imaginary map in front of her. She indicates the location of a *fast food restaurant* with her left index finger while saying “and there’s McDonalds” ("Und da iss McDonalds"; Fricke 2007:128). The non-dominant right hand index finger is a *street* that defines the *street level*.\(^{43}\) The *street level* is the plane on which *buildings* are located and on which the interlocutor is supposed to imagine the described route (ibid., p. 129).

The index finger of the left hand is profiled as a line (Figure 8.14b). The continued line of the left hand index finger intersects with the surface of the *street level*. The intersection of a line and a surface is a point, the geometric entity that conveys no spatial information of shape but merely, and most precisely, a location. Thus, the *fast food restaurant* can be big or small,

---

\(^{42}\) The term ‘Pointing At Object’ does not further specify between the gesture types Pointing At Object (contact) and Pointing At Object (no contact). Thus, both types are meant.

\(^{43}\) The right hand, the *street* that it represents and the *street level* are not part of the systematic form analysis conducted here. Still, this is an example of the gesture form of two separate articulator units (right and left index finger) working together. Further, the *street level* that is shown with the help of the index finger of the right hand is an example of the contribution of context to Gesture Form Analysis: the index finger alone cannot represent a 2D surface, rather, it is conceptualised as something linear. The knowledge that a street is normally located on ground level and that maps are normally laid out horizontally supplies enough context to derive the orientation of the 2D *ground*—even by means of an articulator that is conceptualised as a 1D line.
flat or high, round or square. The gesture informs only about its location in relation to the street on a “map-like space” (Fricke, in print). The direction conveyed by the Continuation of the line the left index finger represents, is just a constituent form that affords marking the resulting form of a point in space by Continuation and Intersection (with a surface). The point, that is the location of the restaurant, is the information to be conveyed in this gesture. The point is the resulting form.

Figure 8.14(a,b): Pointing At Location (no contact) on an imaginary surface. Adapted from Fricke (2007:129; 2014:1818; excerpt). Courtesy of Ellen Fricke (www.ellenfricke.de) and De Gruyter; through Copyright Clearance Center. Copyright © 2007, Walter De Gruyter GmbH.

In an alternative interpretation, the fast food restaurant could be imagined in three dimensions, including its height, on the three-dimensional map and the index finger touches the top of the restaurant with its tip—thus making the Continuation of the profiled line segment unnecessary. Such an interpretation of gesture form will be analysed with the help of another gesture in the following section. Since the object in this gesture is physically present, it indisputably involves contact of the articulator and the object.
8.4.5. Pointing At Object (contact)

The operation that has to be added to Pointing At Location in order to accomplish a prototypical pointing gesture (Pointing At Object) is Adjacency. When we point at something, following Fricke (2009:105), we indicate an object that is adjacent to a specific location. The operations treated in the sections above allow us to draw attention to this location. For pointing at an object instead of just a location, the gesturing body necessarily has to be understood as coupled with the environment (Goodwin 2003b:20), specifically, in the case of Pointing At Object, with the reference object.

Already Pointing At Location necessarily involves an additional entity, at which an intersection of two forms can occur (Intersection always requires two input forms, Section 7.2.3). This entity can be an imagined form (Figure 8.14b) as well as the surface of another articulator (Figure 8.22b) or the surface of another object (Figure 8.12b). As noted above, if we understood the gesture portrayed in Figure 8.14b as pointing to a three-dimensional fast food restaurant, not only to its location, it would thus also have to be reckoned as an example of the type Pointing At Object.

As is the case for Pointing At Location, Pointing At Object also occurs as two related types that are minimally different: one that does and one that does not involve the operation of Continuation. Excluding Continuation, the object is in contact. Including Continuation, the object is not in contact with the articulator. The technical term carries the differentiation between the two related gesture types in parenthesis (no contact/contact). As mentioned earlier and as used repeatedly in this chapter, leaving out the parenthesis conflates both types and allows for propositions that are true for both types.

For the analysis of the stereotypical pointing, we begin with the kind of pointing that is in contact with the referent. Figure 8.16a shows a two-year-old Swedish girl indicating a part of...
the body of her interlocutor that she saw in a picture in a book while saying “eyes” (“ögon”). In contrast to the way most adults would point at the eyes, the girl actually touches the interlocutor’s eyes with his index finger. Her index finger is the profiled articulator, a line, that intersects with the surface of the interlocutor’s left eye or eyelid in a point (Figure 8.16b). Not the location, but more precisely the object adjacent to it, the actual eye, is the referent in this gesture.

Since the object pointed at is small and could be conceptualised as a point itself, the difference between a location pointed at and an object pointed at is not marked in this example. The examples in the next section show objects that are much larger. Again, the very general operation of Adjacency, being open to forms of zero to three dimensions, allows applying this gesture type to entities of very different sizes and shapes. Its adaptability is one of the reasons why this operation often comes last in the order of gesture form operations. It adapts gesture form to spatial and material context and allows the low-dimensional constituent forms in pointing gestures to be compatible with the three-dimensional objects in the spatial surrounding.

Children often actually touch the nearby things they are pointing at. “This means that the typical pointing gesture of children is both contiguous to and directed to its target” (Andrén 2010:184-5). Note the (original) italics emphasising two characteristic aspects that are, in the dissertation at hand, treated as operations of gesture form: Adjacency (in other words: spatial contiguity) and the one-dimensional Shape Profile that indicates direction, as used in Figure 8.6 through Figure 8.8. Since these kinds of pointing need one operation less (Continuation), they may be easier to process. This prediction of the present typology goes hand in hand with
the assumption “that contiguity is the primary factor behind children’s early comprehension of pointing gestures” (Andrén 2010:185).

8.4.6. Pointing At Object (no contact)

One example of the most prototypical way of pointing, Pointing At Object (no contact), is a Zinacantán girl pointing at another person that she wants to pick lice from (in a pretend-play context, Haviland 2003:146, Figure 8.18a). In Figure 8.18b, the profiled articulator (index finger) is profiled as a line that is continued away from the speaker until it intersects with the surface of an object, in this case another person not shown in the original picture. However, not the location of the point itself but rather the person adjacent to this point is referred to in this gesture. Thus, the operation of Adjacency needs to be added.

The question of which articulator pertains to the active articulator unit (operation Articulator Profiling) is relevant for the interpretation of the gesture in Figure 8.18. It may change the indicated direction significantly. Defining precisely which articulators are profiled might seem convoluted, but the girl’s pointing gesture is one of the examples that show a different profiled articulator unit to result in a different pointing direction: If the articulator unit included arm and hand, the small index finger would have to be regarded as deviating from the predominant orientation of the overall articulator unit. Including the arm into the profiled articulator unit may just result in pointing to the chest in contrast to pointing to the head of the targeted person, which both seem to cater to the same communicative goal of identifying who to pick lice from. It is further possible that the height of the gesture is altogether backgrounded as was the case in Figure 8.9b and Figure 8.10b.44 Haviland’s example

44 For an experimental study on pointing gestures that point ‘too high’, see Kranstedt et al. 2006.
and the example in Figure 8.9b and Figure 8.10b suggest that ambiguous Articulator Profiling may occur often in pointing gestures, when both competing profiles allow the successful individualisation of the referent.

The person pointed at is not part of the original illustration, as noted in Figure 8.18b. Illustrations in gesture literature concerned with pointing seldom include the reference object of a pointing gesture, since, in a naturalistic setting, participants are free to point at any object outside of the camera view. Even if the object is part of the photo, the orientation of the body parts articulating the pointing gesture is difficult to map precisely onto the spatial scene because two-dimensional depiction and lens distortions simply do not allow for the three-dimensional overview that subjects in the original scene made use of to perform and interpret pointing. In this sense, two-dimensional illustrations of pointing gestures may show the structure of gesture form, but not for example whether the continuation of the profiled line actually precisely intersects with the referent. Motion-capture studies can fill this methodical gap (Kranstedt et al. 2006), with the drawback of them seldom being conducted in naturalistic settings.

An exception to the lack of pointing referents in visual gesture data is Goodwin’s analysis of pointing gestures at an excavation site (Goodwin 2003a:233). The archaeologist on the left side of Figure 8.19a indicates to her student on the right a “problem area” (ibid.) by a Pointing At Object (no contact). By calling it “a problem area”, it becomes apparent that a two-dimensional form is referred to. This emphasises the need for the operation of Adjacency, since the operation of Adjacency is open to forms of any dimensions and connects the location pointed
at with the two-dimensional area that this pointing gesture refers to. The indicated point illustrated in Figure 8.19b is part of the area that is identifiable by the professor (left) and student (right). The object pointed at is actually a surface area that the point (resulting from the intersection of the continued line and the surface of the ground) lies on.

Figure 8.19(a,b): Pointing At Object (no contact), in this case an area on the ground. Adapted from Goodwin (2003a:233; excerpt). Courtesy of Charles Goodwin and Taylor and Francis Group, through Copyright Clearance Center. Copyright © 2003, Taylor and Francis Group LLC Books.

*Comment on pointing precision*

The gesture just analysed (Figure 8.19) as well as those in Figure 8.16 and Figure 8.12 are pointing gestures with a high degree of precision. This is not always the case. There is always a certain inevitable leeway in interpreting a pointing gesture, especially when the possible targets are far away. The precise spatial information that can be derived from the constituting forms of these conducted analyses should not hide the fact that pointing gestures can also be vague. The form of a line is more of an idealised path to follow by an observer. If the precise indicated direction does not yield a target that is acceptable in the context of the utterance, other objects in the vicinity may be considered. This leeway is not part of a Gesture Form Analysis, but of other studies (e.g., Kranstedt *et al.* 2006) and there are certain cues that indicate, whether the gesturer is putting a lot of effort in pointing precisely (Hassemer 2009:69-71).
The identification of a line as part of the gestures analysed so far, thus, is not to claim that pointing gestures are always precise, but rather to distinguish them from those kinds of pointing, in which the gesturer desires to be vague about the pointing direction. Those pointing gestures for which directional vagueness is a constituting characteristic and the amount of vagueness can be defined by the spread of the hand and the degree of lateral movement (see Section 8.6.1).

**Ambiguity of ‘later’ gesture form operations**

The pointing gesture in Figure 8.19b includes an ambiguity with regard to certain operations of the Gesture Form Analysis. While it is necessary to perceive the rough pointing direction (1D Shape Profile, Continuation), the following two operations (Intersection, Adjacency) might not be necessary, simply because there is no other “problem area” immediately close by to cause confusion about the intended referent object. In the context of a routine conversation, gestural reference to this area might be so common that its interpretation gets reduced to roughly identifying the gesture’s direction. Interpreting the index-finger pointing hand as Directing (Section 8.4.2) could then be sufficient in order to individualise the referent. The two alternative gesture form interpretations, with or without the last two operations of Intersection and Adjacency, are difficult to discriminate; yet, I opt for categorising the gesture in Figure 8.19 in its full potential of Pointing At Object. Pointing At Object starkly contrasts to examples of Directing as in Figure 8.10, which clearly involve neither Intersection nor Adjacency. This contrast is one reason to code Figure 8.19 as including the last two operations. Further, also an observer that neither shares the general archaeological knowledge of the interlocutors in the picture nor the specific experience of working at this specific site in Figure 8.19 (e.g., the reader of this dissertation) is, in general, able to derive the spatial information that discloses the rough location of the “problem area” indicated by the performed gesture. Gesture Form Analysis will therefore not include whether the specific observers in Figure 8.19 might not need to process certain gestural information because they already know where exactly the referent is located (the gesture thus being at least partially redundant). Many sources can afford indicating the referent and thus overlap with the spatial information derived from the gesture. Not only can the observer know the location of the object already: speech can indicate where the object lies, a prior gesture might have precisely indicated the object already, or the object might be so dominant in the surrounding that no indication is necessary. This is consonant with Talmy’s analysis of gesture and speech in that the
communicative situation offers diverse, and possibly redundant, “cues to the target” (Talmy 2012; forthcoming).

There is no reason to exclude the spatial information of the last two operations just because what they convey is redundant to what is indicated by another modality or context. Whether every part of the gesture’s potential to individuate a referent is exploited in a specific case is left to a very complex, in-depth analysis of all modalities and context. For now, it will be assumed that gesturally indicating the referent while also giving other cues that indicate the same referent together afford unambiguous communication. Together, they not only make it possible to individualise a referent, but, by virtue of their redundancy, mutually reaffirm the correct interpretation of verbal and gestural indication to the observer.

**Ambiguity with regard to the operation of Intersection**

With the example presented in Figure 8.19, we touched upon the ambiguity of whether or not some of the ‘later’ operations are to be included. Another kind of ambiguity within certain operations arises, for example, when someone is pointing with the index finger in the direction of a window while simply saying: “There it is!” In such an ambiguous example, the gesture alone does not afford specifying how far the profiled line has to be continued and where exactly the intersection occurs. An observer could search for the referent in different locations. One could suspect the referent (1) anywhere along the imaginary line intersecting with a small entity in the air or transparent surface, like a lady bug on a windowpane, (2) at the Intersection of the line with the surface of the next large object, say the building next door, or (3) where the line intersects with a site that is far away, occluded behind other physical objects, but known to both interlocutors. All these alternative interpretations include the five operations of Pointing At Object (no contact), as illustrated for example in Figure 8.19. They have one and the same structure of gesture form, but they are still ambiguous with regard to the referent. This is not exactly an issue specific to Gesture Form Analysis, but rather a characteristic of pointing gestures in general. Indications selecting one of several possible points of intersection have to be given by other gestures, by speech, or by other context. Talmy’s current work elaborates these so-called “cues”. He distinguishes different categories, one of which is “gestural cues” (Talmy 2012; forthcoming). The kind of ambiguity discussed in this section is an example of the fact that gesture form operations do not constrain the conceptualisation of gesture form in a way that singles out one target unambiguously. Just as in grammar, some cases result in an underspecification that does not, on its own, allow for clear communication.
(Talmy 2000:233-42). In these cases, further context is necessary to negotiate which of the different possible targets is meant.

Besides proposing a basic spatial structure of pointing gestures, Gesture Form Analysis aims at the specification of ambiguities inherent to the structures of gesture form across the discussed gesture types. The last three sub-sections touched upon different kinds of context, especially speech, which are sometimes necessary for disambiguating gesture form. This was also done to sketch some of the limitations of the gestural modality, and to highlight the dependencies of gesture from speech as well as other contextual factors (see also Mittelberg & Waugh 2009; Müller 2010).

8.4.7. Drawing

![Typology of pointing gestures extended to the types of Drawing (profiling the articulator as a 1D shape) in a single tree-diagram.](image)

To arrive at the gesture type Drawing, coming from Pointing at Location, only one operation has to be added: the first and only gesture form operation that specifically interprets motion. To understand, for example, that an imaginary line is being Drawn in the air, conceptual processes that include the dynamic gesture form operation Trace Leaving are necessary. The term ‘dynamic’ marks this operation as one that processes only the change of position of or within a given form. This contrasts to all other operations applied in this chapter, which process static form aspects. All these operations function without reference to any movement during the expressive phase of a gesture. That is, we can walk into a room and see a person Pointing At Object (no contact) without having witnessed any movement, and still understand the gesture. As will be shown in all dynamic gesture examples of this dissertation, ‘static’
gesture form operations also occur in gestures with motion, but they do not involve the interpretation of motion.

With the type Drawing, we cross the border between deictic gestures and iconic gestures; a border that, I would like to claim on the basis of the structure of the form analyses conducted here, is not bigger than the border between other pointing types. We cross this border by simply adding one operation, in the same way that we added one operation to get from one type of pointing to another. Gesture form operations serve to contribute to consolidating the landscape of gesture types in two ways. They can be used to (1) tease apart the pointing types mentioned so far, but also to (2) include them in one group of gestures that profile the articulator as a one-dimensional shape: the articulator is perceived as indicating a direction (the central affordance of one-dimensional shapes, Figure 7.8). Drawing belongs to this group by virtue of the index finger being conceptualised as a line; just as the above-analysed examples, including Directing or Pointing At Object. The structural difference between Drawing and some types of pointing is just one gesture form operation. Drawing is, structurally, at least as similar to pointing gestures, as different pointing types are similar among themselves, while the structural overlap of two or to three operations between pointing gestures and Drawing is comparably large.\(^{45}\)

Drawing, by definition, conceptualises motion in a specific way, more specifically, motion is interpreted as leaving a trace. The output form, a path or curved line, is as large as the movement of the gesture, or proportionally larger for rotational movements. Its shape is defined by the trajectory of the movement. The difference between motion being interpreted as motion and motion interpreted as leaving a trace is apparent throughout the gesture literature, for example in two gestures analysed by Efron (1941) in Figure 8.21 (for detailed accounts of extension deriving from motion, see, e.g., Mandel 1977; Mittelberg 2010a; Taub 2004).

Figure 8.21a shows motion depicted as motion. The size and maybe even the exact form of the hand trajectory might not be mappable onto the actual motion of the person diving into

\(^{45}\) (1) Pointing At Location (no contact) and Pointing At Object (no contact) coincide with Drawing (no contact) in the three operations 1D shape profile, Continuation, and Intersection. (2) Pointing At Location, Pointing At Object coincide with Drawing (contact) in the two operations 1D shape profile, and Intersection. (3) Pointing At Location (contact) and Pointing At Object (contact) overlap with Drawing (no contact) also in the two operations 1D shape profile, and Intersection. This already excludes the overlap in the operation of Articulator Profiling, since as it is always the first operation.
the pool, but they roughly stand for the trajectory of the jumping person. At least the depicted arrow in Figure 8.21a certainly does not show the shape of a man, nor that of a pool. The motion of the hand stands for the motion of a person diving into a pool. In Figure 8.21b, on the contrary, the motion of the hand has to be translated into the extension of the “row” referred to in the accompanying speech. Conceptualising the row to move along this path does not appear to be the correct interpretation, but rather the horizontal (and maybe partially also vertical) extension of the row from left to right (given the sparse verbal context), maybe a row of people or a row of other objects. This latter case, which entails the operation of Trace Leaving, can be called a gestural application of fictive motion (Talmy 2000:99-175; see also Section 3.3) just that in gesture, the fictive motion is expressed not by a word but rather by a moving articulator. Although the articulatory motion is factive, it is fictive in the sense that the entities referred to actually do not move as the depicting finger moves. Moreover, they may not have come to be arranged in the form of a row by the motion performed in Figure 8.21b (e.g., people queue up from left to right) but by an unrelated motion (e.g., people join one person in a centrifugal pattern, by lining up both to her right and to her left).

Figure 8.21: Examples of understanding motion as motion (a) and motion as leaving a trace (b). From Efron (1941:97, 74). Courtesy of Columbia University Press. Copyright © 1941, Columbia University Press.

Drawing seems to be much more frequent in contact than without contact. ‘In contact’ does not mean that the articulator has to be touching a physical surface. It means that the created form, whether there is a physical surface to draw on or not, is to be imagined in contact with the articulator. This might entail imagining a surface that intersects with the line the articulator is conceptualised as, as in drawing on a surface. Alternatively, the surface of the fingertip could also be imagined to intersect with the line. Both interpretations are usually identical in gesture form: a point of intersection that leaves the trace as its imaginary result, or, put more simply a line drawn in the air.

46 This trajectory is simplified into a one-dimensional path. The actual motion pattern requires an array of paths, but this kind of manual depiction regards only the diver’s predominant spatial translation.

47 See Section 3.3 for cases in which conceptualising an external surface is non-intuitive.
In the example of Figure 8.22a, part of the gestural depiction of a round object, the participant emphasises an already-present shape aspect of his left hand (the left hand’s gesture is not being focussed on in this analysis). The circular movement of his right hand index finger on the middle finger and thumb of the left hand conveys the one-dimensional contour of an arc of a circle, reemphasising the circular shape presented by the left hand (Figure 8.22b). The participant chooses a dynamic way of displaying this contour, using the movement of the tip of the index finger as a metaphor for the extension of the depicted shape. As in prior examples, the articulator is conceptualised as a line that intersects with the surface of the other articulator (the left hand) in a point. This point moves, along the left hand middle finger and thumb. Movement here does not stand for the movement of the referent, but is rather translated into the outer contour of the referent. The point leaves a trace in the form of an arc of a circle, in topological terms a one-dimensional path, a bent line. Though the spatial information conveyed in this gesture is far from describing a complete circle, it emphasises the key aspect of this shape (redundant with the left-hand gesture): being regularly curved.

Note that Drawing is characterised by two lines; two one-dimensional constituent forms. First, the articulator is conceptualised as one-dimensional. But this line is not yet the resulting form. Rather, this line intersects with a surface in a point, which, leaving a trace, results in a second line. This line can be of any curvature and even form a circle thus affording the depiction of any path in three-dimensional space, for example the contour of an object.

Figure 8.22(a,b): Drawing (contact) on the surface of the other hand held still (@nmLab).

The gestural type Drawing shares aspects with real physical drawing in that a “pointed implement” (Mandel 1977:67) touches a surface at a point, and by contact, leaves a trace in
form of a path. Actual contact has the same effect as the operation of Intersection (Section 7.2.5), thus following the same pattern. Prototypically, physical drawing implements are line-like or at least narrow towards their tip. They extend in one direction and offer a well-defined, often point-like tip (pencil, pen, fountain pen, chalk) that leaves a linear trace. A non-point-like tip (text marker, big paintbrush, chalk held flat, filter membrane) can afford a trace of higher dimensions, often a surface (these systematic elaborations of Drawing are dealt with in Section 8.7.3).

Figure 8.23a shows the result of the action of inscribing in the sand, an archaeological technique (Goodwin 2003a) analogue to the structure of gesture form in Drawing. The tool can be of any form that touches a surface at a point. The contact point will leave a trace as it is moved, in Goodwin’s example the trace depicts a contour (Figure 8.23b).

Figure 8.23: Physical drawing with an instrument exhibiting similar structure to Drawing (contact). Adapted from Goodwin (2003a:231). Courtesy of Charles Goodwin and Taylor and Francis Group, through Copyright Clearance Center. Copyright © 2003, Taylor and Francis Group LLC Books.

With regard to gesture form, the action of inscribing (Figure 8.23b) is strongly related to gestural Drawing (Figure 8.22b): both characteristic operations and the dimension of the output form are equal. This example is to again highlight the structural similarity of gesture to other practices with a visible, physical effect (see also Streeck 2009). The combination of gestures form operations is rooted in the structural similarity to actions that these gestures replicate, in an abstract fashion. The different dimensions of the constituent forms that make up one gesture are not mental products specific to the conceptualisation of gestures, they are also inherent to the structure of how we act in our physical surrounding—be it by the pointed implement leaving a linear trace in inscribing (0D and 1D), or by the surface contact (2D) of
hand and ball (both 3D) in playing basketball (see discussion of form affordances in Sections 3.6 and 7.1.2).

Another variant of the gesture type Drawing that also highlights the structural similarity to the physical practice of drawing can be found among Goodwin’s examples (2003:21). Here, the archaeology teacher who is squatting left in Figure 8.24a shows her student (on the right) by gesture, along which path to inscribe next. Goodwin indicates the path with a small white arrow drawn onto the screen shot (Figure 8.24). Thereby he unwittingly identifies a core aspect of the gesture type Drawing by his own drawing onto the image.

This gesture does not appear involve touching the ground with the index finger, already because this would actually leave a mark in the sand and the teacher is only instructing the student. Thus, the direction indicated by the finger has to be continued until meeting the ground (Figure 8.24b). As shown in Figure 8.20, Drawing varies in being performed in contact (Figure 8.22b); or without contact, requiring the operation of Continuation (Figure 8.24b).

Figure 8.24(a,b): Drawing (no contact) without touching the ground. Adapted from Goodwin (The Body in Action (2003b:21), Palgrave Macmillan; excerpt). Reproduced with permission of Charles Goodwin and Palgrave Macmillan. This material may not be copied or reproduced without permission from Palgrave Macmillan.

Drawing can be used just as well to draw a line whose shape is to be interpreted metaphorically as mentioned earlier in Müller’s ‘love relationship’ example (Section 3.3; Müller (2008b:233-8), which requires the same operations as in Figure 8.22 or Figure 8.23. For the
perspective of Gesture Form Analysis, iconic and metaphoric gestures are taken to equally build upon gesture form with all its constituent forms.

Further examples of imaginary lines drawn in the air are documented by Mittelberg (2010a), who in investigating university lectures reports on a broad range of drawing gestures used metaphorically to depict the structure of abstract concepts of linguistic theory like sentence structures or word classes, as reproduced in Figure 8.25. She explicitly compares different resulting forms, emphasising the dynamics of “fleeting hand movements that draw [...] imaginary traces in gesture space” (Mittelberg 2010a: 367).

![Figure 8.25: The resulting traces of Drawing gestures. From Mittelberg (2010a:368; excerpt). Courtesy of Irene Mittelberg and Equinox Publishing. Copyright © 2010, Equinox Publishing Ltd.](image)

Semiotic relations as gesture form operations

As introduced in Figure 7.13, gesture form operations can be read as an application of semiotic relations (between the sign and the object) to the spatial modality of gesture. Since this depends heavily on the kind of semiotic relation, the following will touch upon those semiotic relations that can be applied as a purely spatial relation, in short, those semiotic relations that are spatialisable.

Both profiling operations could be called a metonymy in that the whole visible body is reduced to a part of it in a totum-pro-parte relationship; in Articulator Profiling a portion of the whole body and in Shape Profiling a shape aspect of the profiled articulator unit. In its similarity to the simple geometric form of a line, Shape Profiling could also be seen to include a metaphor, since the physical articulator stands for something abstract, the line.

Continuation could then be considered as the reverse metonymic process, a pars-pro-toto relationship, in that a line segment stands for an indefinite line. While the part-whole relationships are apt to group the operations as done in Figure 7.14b, they cannot specify the relation of constituent forms that structure pointing and drawing gestures, as illustrated for example in Figure 8.20.
The similarity of gesture form operations to indexicality and directionality might be even more salient, for example in pointing gestures, such as the ones discussed in this chapter. One could identify indexicality and directionality already in the profiling of a line, since it is defined in its precise rotation, but they could just as well be attributed to the continuation of the line, since with the one-sided continuation of a line, a direction is given.

The operation of Trace Leaving can be seen as a metaphor, but very different from other metaphors that are open to subjective and rich domains, such as ANGER IS HEAT IN A CONTAINER. Trace Leaving is a specific kind of metaphor in that it is readily spatialisable. Exactly the space travelled by a certain form during a certain movement (input form) becomes the trace (output form) that is conceptualised in Drawing. Movement is a metaphor for extension (EXTENSION IS MOVEMENT). To be open to mathematical approaches and motion-capture methods within Gesture studies, this formalisable operation of a clear input and output form is actually not a metaphor, but merely a simple geometric operation.

Overall, the Gesture Form Analysis of the group of pointing and drawing gestures could not have been described in their specific spatial constraints with semiotic relations. Gesture form operations allow for a higher structural granularity of analysis—at least with regard to form—from which overlaps and differences of gesture types are discernible.

### 8.5. Surface Contact Gestures: Holding and Moulding

The second group of gesture types is defined by the articulator unit being profiled as a surface. Often, this surface is the tactile surface of the hand: the palm and the inner side of the fingers from the palm up to and including the pads. When only single fingers or the thumb are extended, they do so predominantly in one direction and are thus often profiled as one-dimensional. By contrast, the whole hand does not only saliently extend in one direction (from the root to the finger tips) but also laterally from thumb to the opposite side of the palm and the pinkie, especially when spread. Thus it is often conceptualised or, more specifically, profiled as two-dimensional. This is true for the large group of gestures that imitate touching and handling objects.

Especially the inner side of the hand is used for controlled interaction with objects. In object interaction, the surface of the hand is adjacent to the surface of the object, sometimes enclosing it. The surface of the hand and the surface of the object are in contact. Put simply, the hands touch the object; they are in surface contact with it. Therefore the group of types will be called Surface-Contact Gestures.
As introduced in Chapter 7, two kinds of surface contact distinguish these gestures. In a gesture that pretends to hold something, the surfaces of the articulator and the surface of the object do not move in relation to each other. If we hold a ball, even when we move it around, any point of the inner surface of the hand will remain in contact with their corresponding point on the surface of the held object (static contact). Gestures that imitate such actions are called Holding (Figure 8.26).

If, on the other hand, we use our hand to mould an object or to explore the contours of an object, our hand’s surface can move in relation to the object while remaining in contact with it (dynamic contact). The larger the object, the more we have to move our hands in order to capture its extent and shape. These gestures will be called Moulding (Figure 8.26).

<table>
<thead>
<tr>
<th>tactile 'surfaces' of the hands enclosing an imaginary object</th>
<th>plus: hands moving to show the shape of the imaginary object</th>
</tr>
</thead>
<tbody>
<tr>
<td>&gt; holding</td>
<td>&gt; moulding</td>
</tr>
</tbody>
</table>

Figure 8.26: Holding and Moulding in a simple form-based differentiation.

As in the preceding section, I will break down these types into gesture form operations beginning with the simplest combination of operations that lays the groundwork for the more complex but very frequent Surface-Contact Gestures.

8.5.1. *The basis for Holding and Moulding: Representing in two dimensions*

As in one-dimensional Shape Profiling, different gesture types use the result of Shape Profiling as a shape to build on by adding other operations. Representing in two dimensions seldom is the resulting form. One example of this rather rare kind is a participant describing a flat oval shape (for details of all objects, see Appendix 5) by referring to her flat right hand that resembles this object in height and width—but not in its thickness (Figure 8.27b). She says that the object is “oval like my hand but more continuous” ("oval wie meine Hand nur kontinuierlicher") and uses her left hand (not part of the analysis) to point to the right hand whose shape is in focus of the gestural presentation.
Her hand is not thin enough (10-20mm) to represent the object thickness (3mm). She also supports the interpretation of a two-dimensional profile by explicitly and separately addressing the thickness of the object right before this gesture, stating that it is very thin and performing a separate gesture to exhibit the thickness. The information from the ‘thickness gesture’ on one hand and the information conveyed by Representing in two dimensions on the other hand, together, fully describe the proportions of the object.

In contrast to the gesture types in the two following sections, the two-dimensional shape profile cannot only be abstracted from the inner surface of the hand (in Figure 8.27b), but also from a two-dimensional version of the hand itself (somewhat in the middle of the hand’s thickness). Both surfaces, at the hand’s surface or running through the hand, result in the same profiled surface, only that this surface is located at slightly different depths in space. The gesture type Representing in two dimensions is not necessarily related to the spatial structure of the act of holding or moulding something, in which the object is always outside and adjacent to the tactile surface of the hand. The following sections will deal with the very frequent group of gestures that require the surface of the hand to be profiled, since the Held\textsuperscript{48} or Moulded |object| is to be imagined in contact with the hand. Since the interpretation of the profiled two-dimensional shape being located at the hand’s surface recurs in multiple gesture types and in multiple gesture tokens analysed in this dissertation, this interpretation will be favoured, as done in Figure 8.27b.

\textsuperscript{48} Capitalisation indicates the gesture type Holding in contrast to the act of holding.
8.5.2. Holding

When we hold an object, the body part doing the holding has to be somehow in contact with the object. The type Holding takes advantage of this fact, so that understanding that there is an object in surface contact with the hand makes the operations of Adjacency, more often specifically Division and Enclosure, a necessary condition for understanding this gesture type.

Static Holding

Within the gesture type of Holding, two kinds are distinguished, one in which the articulator moves, dynamic Holding, and one in which the hand doesn’t move, static Holding. These two kinds are not distinct gesture types, they are just distinguished so as to mark that Holding can also contain motion, but motion does not stand for something else, as it is the case in Moulding.

Some Surface-Contact Gestures, such as the ones performed with “palm open, fingers extended more or less loosely, palm turned upwards” (Müller 2004:233), allow assuming an object being placed on top of the inner surface of the hand. In these gestures, the hand resembles a hand physically supporting something (Müller 2004; Kendon 1995). In Figure 8.29(a), a participant (introductory questions, Study 1) reports on the things that are currently on her mind and explains that in case her failing the exam, she “would have done the French course totally in vain” (“dann hätte ich halt Französisch total umsonst gemacht”). She raises her left hand rotating it to present the hand palm up for an instant, while uttering the word “French”. Then she immediately retracts her hand again. The whole movement is quick and lax. This gesture resembles the so called ‘palm up open hand’ gesture (e.g., Müller 2004:241), in which the speaker “presents the discursive object” (ibid., p. 242), seeking for agreement with the uttered proposition or expecting the interlocutor to signal his understanding or ask a follow-up question. In this case, the object could stand for the French course (in a narrow interpretation) or the negative effect of her not receiving credits for the French course (in a broader interpretation).
The ‘palm up open hand’ is an example for giving only minimal information concerning the |object|’s shape, since the hand shape is not necessarily adapted to the shape of the |object|. The hand’s lax curvature presented only some milliseconds does not seem to transport form specifics of an imaginary object. It could thus Hold a |small marble| as well as a |larger box|. Figure 8.29(b) introduces the graphical code of marking a cube with “ANY” to stand for any shape of up to three dimensions. In this example, hand’s profiled surface just roughly indicates the location of the |object| (on top of the hand and adjacent to it), but it does not necessarily convey anything of the object’s shape.

If it was given by the context that the shape of the hand was to exactly delimit the |object|’s lower side, the last operation would have to be Division instead of Adjacency, because the surface of the hand precisely divides the space occupied by the visible hand from the space occupied by the invisible |object|. But since the rather lax hand shape is presented only very briefly and the |object| here stands metaphorically for language course or for a consequence of a her theoretical failing in an exam, Figure 8.29b visualises the application of the general operation of Adjacency, leaving the object’s shape and dimensionality undefined. Since the physical act of holding generally involves physical, three-dimensional objects, the gesture’s |object| may most likely be interpreted as being three-dimensional, but the possibility of

49 See also Mittelberg (2010a; more examples and semiotic analyses, ibid., pp. 361, 365).
conceptualising this discursive object as, for example, a point-like entity cannot be excluded. Mittelberg also distinguishes relaxed hands from more precisely defined hand shapes, which allow to infer the shape information of the referent when “the hand shape is more salient than the contextual movements” (Mittelberg 2010a:355).

Also in two-handed gestures, knowing that something that we cannot see (the |object|) is supposed to be imagined adjacent to something that we can see (the hand), does not automatically allow us to infer a precise shape. Still, in two hands limiting the |object| on two sides, one extent of the |object| is revealed. Adjacency of two planes that face each other, partially enclosing an |object|, specifies the scale of the object on one axis (see, e.g., Figure 8.30b).

Adjacency can be vague with respect to the output form (Adjacency in Section 8.4). In Holding, the object can be specified on one or more axes, but also in its entire shape. When an articulator’s shape conveys rich spatial information, as in precisely displaying a curved surface, it can enclose an |object| and thus large parts of the |object|’s shape can be communicated. Enclosure is the most specific subcategory of Adjacency. It includes Adjacency but is more constrained in the dimension of the output form (Chapter 7.2). Surface Contact Gestures include everything from general Adjacency, through Division to partial and full Enclosure (see also Levinson’s & Wilkins’ “inclusion”, 2006:522).

In simply Holding an |object| between two not precisely defined hand shapes, instead of just one, the object is only minimally enclosed. In McNeill’s classic example (Figure 8.30a, “it was a Silvester and Tweety cartoon”, McNeill 1992:14-5), the metaphorical object that stands for the genre of Tweety and Silvester cartoons can be of very different shapes and its precise shape does not seem to matter. Still, the two hands enclose it and limit the |object|’s size. To obtain this information, both hands are interpreted as the profiled articulator unit, of which again the inner surfaces are profiled. These surfaces allow us to estimate a distance, the only precise obtainable information about the |object|’s proportions. The |object| can be shaped variously, but we know that its extension on the horizontal axis will be that of the distance between the inner sides of the articulators (Figure 8.30b; for more examples of this kind, see Mittelberg 2010a:353,363,364). As already noted, the width of the |object| does not have to play a role in the overall utterance as is the case here (Figure 8.30b and Figure 8.31b), but it can—a gesture like the ‘cartoon gesture’ does afford showing size on one axis. An example of this would be the same gesture accompanied by “the fish was this big”, or a structurally similar gesture depicting the height of an object (Figure 8.32b).
The |object| elicited in Holding is not only defined in some of its proportions but in its position (as is the case basically in all gesture types, not only pointing gestures). While the information of location is often backgrounded, it is central in the case of a professor depicting a subcategory of auxiliary verbs by gesturally spatialising the hierarchy of word classes (Figure 8.31a). In Mittelberg’s analyses, contiguity relations, such as adjacency, are described as the interface between the visible articulator and the imaginary, but adjacent |object| (Mittelberg 2008; 2010a,b, 2013; see also Section 2.4).
As well as a metaphorical object, a physical object can be referred to in a similar way. In Figure 8.32a, the participant shows the size of an object “in its thickness” (“in der Dicke”). There are two differences in gestural presentations between the example in Figure 8.31b and the following gesture example in Figure 8.32b, which demonstrates the flexibility of the structure of gesture form at the example of the type Holding:

In Figure 8.31b, (1) both hands, specifically their inner surfaces (except for the ring finger and pinkie holding the chalk), are profiled to create a three-dimensional form, (2) which is defined on the horizontal axis only. This form is to be mapped onto the abstract category of a word class. In the following gesture (Figure 8.32b), (1) only the index finger and the thumb are profiled, partially enclosing a three-dimensional form which is also (2) defined on one axis only, but this time, in its height. Figure 8.31b and Figure 8.32b show the structure of gesture form to be (1) articulator-independent and (2) orientation-dependent. A Holding gesture may be performed with the hands or with the fingers and it may convey the size of the object on the horizontal axis or on the vertical axis.

Figure 8.32b shows further that a Holding gesture of this kind can be conceptualised in one-dimensional (vertical) space without losing any of the specified spatial information. This gesture is part of the object description, in which the one-dimensional height conveys just one extent (on a vertical axis) of a physical, three-dimensional, reference object. The specific group of Holding gestures performed by the index finger and the thumb exemplified in Figure 8.32b will be called Measure Gestures (see also Streeck (2008:292) on “bounding” gestures) and is investigated more closely in Part III.
These two examples also mark another core characteristic of gesture form. Gesture form is independent of the physical existence of a referent: the form created in gesturing can be to describe a physical object as well as to describe an abstract concept. As shown in the prior sections, gesture form is also a necessary part of different ways of pointing to physical and imaginary entities. The analyses in Figure 8.30 through Figure 8.32 also show that the same structure of gesture form occurs in a variety of sizes and in different orientations, as also identified by Mittelberg’s (2010a:363) application of the semiotic relation of (outer) contiguity to gestures of this form.

As introduced in 7.2.1, the order of gesture form operations, as the one indicated in Figure 8.32b, is part of a Gesture Form Analysis. Any constituent form depends strongly on all the forms to its left (the last row in the coding scheme at the bottom of the figures). A form feature (in this case the tactile surfaces) can only be profiled when the articulator to which this form pertains is already profiled; and the object including its one-dimensional specification can only be obtained when the two surfaces have already been profiled.50

There are also gestures that convey more specific shape information.51 In contrast to Figure 8.32, Figure 8.33 shows the participant’s profiled articulator unit completely enclosing the imaginary object, thus conveying the shape of the |object|’s complete surface. The complete enclosure is also the reason why the profiled shape is difficult to illustrate. The hands completely occlude the hands’ inner surface as well as the |object|. The resulting form of gestures that exhibit complete, or close to complete Enclosure (Figure 8.33, Figure 8.36 and Figure 8.41; Mittelberg 2010a:362) is often difficult to depict, since the |object| is occluded by the articulator. These four examples support a praxeological view on gestures (Streeck 2009): it is not only the observer’s visual access to the |object| that is relevant (sometimes we cannot see what is being depicted). It is also important that the observer him- or herself imagines the performance of the physical action and partly relives the experience of interacting with an object in a certain way, thus inferring the |object|’s shape (as in the example of an |object| that is Held in a way that visually completely occludes is, see Figure 8.33b).

50 Nevertheless, the configuration that is used to profile the opposing surfaces will reaffirm the (earlier) choice of the articulator unit, and the one-dimensional specification of the referent can also be roughly estimated without the profile of the tactile surfaces of the articulator.

51 The gestures in Study 2 deal specifically with the description of objects’ shapes.
Generally, an alternative to a Holding interpretation of gestures like the one in Figure 8.33 is the hands themselves three-dimensionally representing the object (Section 8.6), which generally occurs less often in the investigated data. For Figure 8.33b, this alternative interpretation can be excluded since the participant clarifies that the object is so small in relation to his hands that it “[...] could have been placed between the palms” (“[...] zwischen die Handflächen hätte nehmen können”).

Since not the hands themselves but what the hands enclose is the referent here, the shape conveyed in Figure 8.33b is three-dimensional only because we know from the context that the referent is a three-dimensional object (object description task). While the resulting form of a gesture as the one in Figure 8.33a can be a solid sphere standing, for example, for a billiard or pétanque ball, it can equally well be a hollow sphere standing for an air-filled balloon or an abstract, entirely empty spherical space. In this case the Gesture Form Analysis would already end with the resulting form of a curved “surface” (third form in Figure 8.33b). In topological terms, two-dimensional is the correct classification for this form, the hollow sphere represented in green. This does not dispute that, in most circumstances, we infer a three-

---

52 The hand configuration is maintained during seven seconds of discourse (“Es handelte sich dabei um einen Gegenstand den man von seiner Größe her ... in etwa ... wenn man die Hände so formt eh em in die ... zwischen die Handflächen hätte nehmen können”)
dimensional object. But additional spatial processing is needed—the operation of Enclosure—to conclude that it is not the profiled two-dimensional shape itself, but the shape enclosed inside of it that represents the object’s shape. As stated more generally in the preceding chapter, two-dimensional shapes afford Division, and, if closed to a hollow sphere, even Enclosure of three-dimensional space (Section 7.2.7-8).

Figure 8.34 shows another case of partial Enclosure. This time, index finger and thumb are profiled as a regularly curved ‘C’-shape, but the exhibited curvature is not spherical as in Figure 8.33b, but only circular or minimally cylindrical (if one was to consider the thickness of the fingers as the depth of a cylinder). Thus the resulting form is not completely defined. The referred object here is a styrofoam sphere of six centimeters diameter, but the spherical (not only circular) curvature is not conveyed in this gesture, but in speech only (“spherical object”, “sphärisches Objekt”). A naïve observer could not conclude anything about the depth of the object. This is an example of less complete Enclosure than in Figure 8.33b, but more complete than in Figure 8.30b and Figure 8.32b.

Figure 8.34: Holding by the complete inner surface of index and thumb, specifying an object’s 2D projection.

The operation of Continuation is coded in Figure 8.34 (but not in Figure 8.32b) because a circle is indicated by the regular curvature of the index and thumb that match a circular form, but the circular form is not complete (but rather only a “C”-shape) without continuing the exhibited curvature to a full circle. Continuation is not coded in Figure 8.32b because Continuation is not a necessary condition for perceiving the height of an imaginary object.
Figure 8.35 (picture *1) shows Calbris’ analysis of a gesture as the one shown in Figure 8.34. In her words, a “dynamic representation” can create similar forms (compare left and right side of the figure). This is consonant with the approach at hand in that Calbris’ stylised graphical illustration shows that both static configurations (*1 and *2, in this dissertation called Holding) and dynamitic tracings (*3 and *4, in this dissertation called Drawing) afford creating the shape of a circular contour.

![Diagram of figures 8.35](image)

A gesture that involves Holding an |object| as in Calbris’s examples *1 and *2 allows different kinds of interactions of the hand and the |object|. One kind of interaction is moving the |object|, as analysed in the following (Figure 8.36b). The spatial manipulation of an |object| is possible for Holding gestures (*1 and *2) but not for Drawing gestures (*3 and *4).

**Dynamic Holding**

The examples of Holding up to this point have not involved motion. The examples of dynamic Holding (Holding and moving) in this section will contain movement as part of the gestural depiction. In Holding, the movement does not influences the shape of the |object|. In actual manipulation of rigid objects, not the shape but the position (or orientation) of the grasped object changes with the movement of body part grasping it. The same is true for gestures of Holding an |object|. The imaginary object is passively being moved with the active articulator. The trajectory of the articulator movement can be mapped onto the motion of the |object|.

Figure 8.36a shows two blended screenshots of a participant while explaining her way through the building up to the lab on the 4th floor (“1” represents the starting position of the
movement, and “2” the end point). She describes how she opened a door by pulling on a vertical bar (“pulled the door open not pushed”, “hab die Tür aufgezogen und nicht geschürzt”). She positions her hand as if grasping an elongated object. The interlocutor is to detect this movement’s similarity with actually pulling a vertical object. Since pulling an elongated object involves arranging the hand around an object, the gesture, conversely, reveals the |object|’s shape by the shape of the hand’s inner surface (Figure 8.36b). Although the inside of this hand shape is not clearly visible to an observer, we still know from our own experience that the hand can accommodate an |elongated object| whose shape is not restricted in vertical extension.

The movement of the hand from the front diagonally backwards is to be understood as incorporating the motion of the |object| (Figure 8.36a). This gesture’s coupled motion results in the |bar| (and maybe also the associated |door|) being pulled open. The gestural motion stands for the physical motion when she opened the door (compare to examples of Trace Leaving, e.g., Section 8.4.7): no additional gesture form operation is needed to understand the spatial relations between hand and |object|. The motion that we see in the gesture is proportional to the motion that we are to imagine the |object| is undergoing, and the gestural movements in the lab can be mapped onto the depicted scene (other examples: Liddell 2003:265-8). This is why the analysis of gesture form in Figure 8.36b is, in its combination of operations, not different from the prior holding examples in Figure 8.30 through Figure 8.33.

Figure 8.36(a,b): Pulling movement in a dynamic Holding gesture miming to open a door (@nmLab).
Metaphoric objects can be moved and modified by Holding and moving, too. Also, different aspects of a movement might be foregrounded, for example the direction of motion in relation to the body, as in an example reported by Müller (2010:167). Participants were being asked to perform a gesture while saying: “He unpacked”, in the sense of revealing confidential information. In response to this task, participants performed a movement, in which confidential information is understood to be “unpacked” by a movement away from the body (Müller 2010:65). The manipulatory action of ‘unpacking’, I would argue, involves moving metaphorical objects from a position close to the body, shielded by the hands, to a position further away from the body, being presented on the open palms (for details and illustration see Müller 2010:65-7). This is just to mention that, in general, the |objects| in Holding can stand for physical objects as prevailing for the object-description task of this dissertation, but they can just as well stand for abstract concepts. Both “iconics” and “metaphorics” (McNeill 1992) can fall into the form-based categories of Holding (this section) or Moulding (next section).

8.5.3. Moulding

![Figure 8.37: Typology of Surface Contact Gestures (profiling the articulator as a 2D shape) in form of a tree-diagram.](image)

As Figure 8.37 illustrates, the gesture type Moulding, an adaptation of Müller’s “the hand moulds” ("die Hand modelliert", Müller 1998:117), builds on Holding. In contrast to the successive addition of operations in pointing and Drawing gestures, the operation of Trace Leaving is added immediately following the operation of Shape Profiling, preceding the operation of Adjacency.

---

53 There was also a trial with a non-metaphorical interpretation of this phrase.

54 A tree-diagram is not able to show both hierarchy and order of operations. In Moulding a trace is left first, and then the adjacent object is inferred.
Unlike in Holding, in Moulding motion is obligatory; more specifically motion gliding over the object. But also unlike with ‘Holding and moving’, the motion in Moulding is not just to be conceptualised as motion. Motion is the source domain for a metaphoric mapping. The metaphor does not work on the resulting gesture form, as for example in the palm-up presentation (Figure 8.29b), but the metaphor is an (early) part of an observer’s understanding of gesture form itself.

Travelled space is to be mapped onto the [extension of an object], just as in Drawing. “The hand moves along a path that depicts the extent and shape of the entity being described” (Liddell 2003:264). Part of understanding both Moulding and Drawing is to interpret one shape to leave a trace; that is, to involve the gesture form operation of Trace Leaving. Among the different ways of how shapes can leave a trace, Moulding requires a profiled surface and co-planar movement. As cited already in Section 2.4, a "surface implement […] can leave a trace shaped like itself, when moved parallel to itself—flat hand moved flat, curved hand moved as if along a curved surface” (Mandel 1977:67).

Trace Leaving and other operations are implied in gesture analyses from a computer science perspective, for example by Sowa in an approach that is consonant with Gesture Form Analysis. In preparing computational recognition of different gestures, Sowa recognised two aspects of form as discriminating features that are also pivotal for the approach at hand. He distinguished gesture types by dimensions and by different areas of the hands being profiled (Sowa & Wachsmuth 2005; Sowa 2006). However, with regard to the specific gesture types, his approach deviates from Gesture Form Analysis. His categories of “dimensional gestures” (2006:253-66) attest only one number of dimensions to each gesture type, thus not differentiating the interaction of several imaginary forms in gesture conceptualisation. Still, the specific calculations of gesture form on the basis of articulator form (although not using these terms) are both precise (to the degree that I can comprehend his extensive computational work) and consistent with the concept of gesture form.

Sowa’s example reproduced in Figure 8.38b (recurrent gesture in virtual-reality route descriptions), is to show the similarities and dissimilarities of his approach to the one presented in this dissertation. There is agreement that the form that results from the movement Figure 8.38 is three-dimensional, a cylinder. Yet Gesture Form Analysis, for its

---

55 In Gesture Form Analysis, every constituent form is defined in its dimension and a gesture type is defined by all the numbers of dimensions of each of its constituent forms.
typology, explicitly distinguishes between the dimensions of the constituent forms that lead up to the three-dimensional cylinder. From the hands, the inner surfaces are profiled. The profiled shapes are bent in its horizontal extent, but straight in their vertical extent. As the movement is performed vertically and straight, it is parallel (co-planar) to the profiled shape, thus maintaining two-dimensionality (if the movement were not parallel, the output form would be of one additional dimension, see Sections 8.7.3-5). Finally, the resulting form, partially enclosed by the curved surfaces, is a three-dimensional cylinder. The next examples (Figure 8.39 and Figure 8.41) will analyse a case of parallel curved motion.

As in the Holding example of Figure 8.36b, only the horizontal extent of the |object| is enclosed by this gesture (Figure 8.38). In these kind of gestures, the actual object can have the height of the length of the motion trajectory plus the size of the hand—or be infinitely high. As with all other gestures, this gestural performance can occur in a context in which certain spatial information on the referent is not relevant or in which approximate information is sufficient. The information could also be not necessary, because it is already known from an earlier gesture or accompanying speech. Gesture Form Analysis focusses on what information can be derived from gestural action. In contrast to Figure 8.36b, the surface that is the output form of the operation of Trace Leaving in Figure 8.38 now encloses a larger part of a solid cylinder, larger than the hand could capture without being moved.

![Figure 8.38: Moulding a cylindrical form. Adapted from Sowa (2006:261; excerpt). Courtesy of Timo Sowa.](image-url)

56 If, on the other hand the movement were horizontal, it would have to be curved and the hands would have to rotate, in order to leave a two-dimensional trace.

57 Due to difficulties in visibility the constituent forms are not depicted in this an the following analyses. Since there is a striking resemblance with simple geometrical forms (cylinder) or with the real objects (Figure 8.39), this did not seem necessary.
In a complex gestural movement consisting of two gestures (a-b and c-e), the participant in Figure 8.39 moulds the Drop (Figure 8.40a). The initial configuration of her hands, on its own (a), would be classified as Holding, but since the hands are then moved and the movement is to be understood as leaving a trace, the participant Moulds an object. When considering only pictures (a) and (b) of Figure 8.39, which represents the first stroke in this gesture sequence, the complete inner surface of both hands is profiled and the depicted object could be a sphere. However in (c), the participant changes the profiled shape of her left hand to only the surface at the pads of the fingers to Mould the slim upper part of the object, while still Holding the object at the bottom with the right hand. This complex gestural movement shows the synchronous use of Holding and Moulding and its effect of conveying large parts of the surface of the referred object.

In general, there are a lot of Moulding gestures in Study 1 and in Study 2, presumably because the stimuli were 3D and Moulding affords conveying shape information on large parts of the objects and can adapt to very different shapes. The important role of surfaces in perception (Section 4.2) supports the frequent occurrence of this category. In Figure 8.41, another participant Moulds the Torus (5.5 cm in diameter, Figure 8.40b), with two hands (in larger-than-life size).
Figure 8.40: Stimulus objects (a) Drop and (b) Torus.

Figure 8.41: Moulding the Torus. The Adjacency is positioned last or second but last (white arrow; @nmLab).

Figure 8.41 exhibits a case in which two alternative interpretations create one and the same form. The operations of parallel Trace Leaving and Adjacency (Enclosure) can change places without changing the resulting form. On the one hand, Moulding the torus could be conceptualised by the inner side of the hand leaving a trace in the form of a hollow torus, which is then filled (by Enclosure) to become a three-dimensional, solid torus (as coded in Figure 8.41). On the other hand, Moulding the same torus could be equally well conceptualised as the hands’ surfaces first enclosing two sections of a curved three-dimensional tube, which, as the last step, leaves a trace of a three-dimensional torus. Since the resulting form is equal in both interpretations, there is no need to distinguish between them. This gesture exemplifies the exception of two operations being commutative (exception to equation (7), Section 7.2.1). As a consequence, both combinatory possibilities are treated as one and the same gesture type.

Figure 8.39 and Figure 8.41 exemplify rich surface information being conveyed all at once, whereas an earlier case (Figure 8.30b) exemplified more reduced form, which can serve as part of a step-by-step description of the different aspects of an object’s shape. In both cases—with
regard to gesture form—surface contact is core for understanding the gesture since it can convey information about the shape of any three-dimensional object, in various ways. Adding to the approaches (Müller 1998; Sowa 2006) that attribute the creation of three-dimensional forms to those gestures that were here coded as Holding and Moulding gestures, Gesture Form Analysis highlights the important role of intermediate two-dimensional forms, at multiple examples of varying complexity (Figure 8.27 through Figure 8.41).

An example showing a necessary reduction of conceptual space

Also among the gesture type Moulding, we find gestures in which space is conceptualised as less than three-dimensional. Another participant describing the Drop uses flat hands to describe the slim upper part of the object including its curvature. The upper part of the object is concave in its vertical extension, but convex on a horizontal plane (when the object is oriented with the peak pointing up as shown in Figure 8.40a). The open hands match this form in that they Mould by bending vertically concave, but afford depicting the object’s shape only partially on a horizontal plane (Figure 8.42). More precisely, with fingers pointing upward, the hands are curved backwards (on the vertical axis in Figure 8.42), from the base of the palm to the finger tips to match the vertical concavity of the object, but they are not simultaneously curved horizontally, from thumb to pinkie to match the object’s radial convexity. No matter at what height we cut horizontally through the object, we would get a circle, a curved form.

However, if the object really was shaped as the inner surface of the hand could be taken to suggest, it would be a (concave) pyramid with flat tapered sides—but this would not match the actual referent (Drop). The Drop has a concave cone as its top part, with a round circumference.

Thus, the hands are to be taken as projecting lines on an imaginary vertical facing plane in front of the gesturer, building an outline in two-dimensional space (Figure 8.42, see also Fricke 2007:267-8 for ‘screen-like’ (“bildschirmähnlich”) reduction of space). The whole gesture sequence is actually the combination of two of these outlines. Figure 8.43 shows the gesture that follows immediately and creates the same shape projection on the mid-sagittal plane. The movement in Figure 8.43 (to be conceptualised as co-linear trace leaving, see pink traces) is performed with an unconventional arm and hand configuration, supporting the claim that the gesture form of these two gestures is to be understood as limited to the bent curves on two

---

58 The accompanying speech consists of the non-complete phrases “yeah on top like-this in its form tapered peak” (“ja oben die so in der form zulaufende Spitze”).
perpendicular planes: a second, separate movement is executed to show the depth contour of the object, forcing the participant to move especially the right hand in a rather uncomfortable way.

Gesture conceptualisation in two-dimensional space was optional in many of the gestures analyse until now. Earlier examples (e.g., Figure 8.22, Figure 8.25 and Figure 8.34) contained resulting forms that can be reduced to two dimensions on a vertical plane in front of and facing the gesturer. The depth of space was not specified and backgrounded in these examples. The gestures in Figure 8.42 and Figure 8.43 were chosen to exemplify a necessary reduction to two-dimensional space as part of gesture conceptualisation. Not reducing one of those gestures to two dimensions would convey a ‘wrong’ object shape (not matching the stimulus). Moreover, the two gestures show the depiction of one the same shape aspect by two consecutive gestures. Only both gestures in combination afford partially depicting the curvature of the tip of the Drop in three-dimensional space.

Figure 8.42: Moulding a conic shape in 2D space (on vertical facing plane) (@nmLab).
Differences in profiled surfaces as part of Moulding

Figure 8.44 shows a very simple kind of Moulding. In explaining that the stimulus object was “flat” (“flach”), the participant moves his flat hand parallel (co-planar) to its vertical extension, thus additionally prolonging the (flat) plane profiled by the hand shape. Although the flat hand shape alone would have already conveyed flatness, the participant emphasises this aspect by performing a movement that extends the established form. This would not have been necessary to encompass the flat area of the object because the object is actually hand-sized, but the partially redundant form information given by the movement (to be understood as leaving a trace) conveys flatness even more saliently. We find similar examples in pointing, when the articulator is moved in the direction pointed at. This does not give new information because the direction of the additional movement is redundant to the finger’s extension, but emphasises the direction already given by the finger orientation.
Figure 8.44 also showcases another characteristic of two-dimensional shape profiling. It is not clear whether, the hand being flat in itself (possessing comparably little vertical extension), the entire hand is conceptually simplified to a plane running through its center, or whether the hand movement is to be understood as imitating someone running their hand over a flat object, favouring the profiling of the hand’s lower surface. Neither of the two possibilities can be excluded. However, the two interpretations exemplify that certain competing resulting forms (plane at the height of the centre of the hand vs. plane at the height of the lower surface of the hand) are similar enough to equally afford conveying the essential spatial information: flatness. Figure 8.32b, in which a distance in the range of centimetres is to be measured precisely between the surfaces of the pads of index and thumb, contrasts to the gesture in Figure 8.44 that conveys the general notion of flatness. Accordingly, some participants (in feedback conversations the official trials of Study 2) were confident in replying that the distance to be measured in an index-finger-and-thumb-gesture as the one in Figure 8.32 has to be computed from the distance between the fingers’ surfaces, not between the centres of the fingers as we would expect for conceptualising the entire fingers as a surface. In Figure 8.44, the precise height of the flat surface is not important making a distinction between these two alternative interpretations pointless. In the absence of cases, in which the profiled surface has to be profiled in the centre of the articulator and given that there are case, in which the profiled surfaces are clearly located at the surface of the articulator (Figure 8.32), I suggest to assume surfaces to be profiled at the surface of the articulator, if there are no clear indications for the opposite.

Returning to the analysis of Figure 8.44, an imaginary object is part of conceptualising this gesture, just as in palm-up-open-hand gestures (e.g., Figure 8.29b). The particularity of the
‘flat’ gesture is that flat |object| is implied to be in surface contact with the plane profiled in this movement. While in the classical palm-up-open-hand gesture the |object| can be of any shape, thus possibly touching the profiled surface only in a point, the object in Figure 8.44 needs to be considered as touching larger parts of the profiled plane—else this gesture would not convey flatness. This makes this gesture clearly involving not the general operation of Adjacency, but the more specific operation of Division, dividing the two solid shapes, that of the hand and that of the |object|, by a plane that is in surface contact with both entities; that is, a plane touching them at an entire area of its shape, not only in a point or along a line. The otherwise communicatively redundant movement thus indeed have a function. It emphasises the application of the more specific operation of Division by highlighting the planar shape that is in surface contact with both the articulator and the |object|.

8.6. Representing in three dimensions

Representing in three dimensions means the articulator stands for something in its entire three-dimensional form. It is different to Representing in one or two dimensions because Representing in three dimensions permits depicting the entire richness of a three-dimensional shape including the representation of complex motion, by re-performing it. No further operations are needed to derive the resulting form so that Representing in three dimensions is just one gesture type. Within this gesture type, two different kinds of referents are distinguished, which make up the next two sections: representing itself (e.g., the hand representing a hand) and representing something else (e.g., a finger representing a standing person). The third section (8.6.3) deals with a very specific group of gestures, emblems, which, in order to be recognised as such, have to be performed in the correct configuration and motion in all three dimensions.

8.6.1. An articulator representing itself in three dimensions

Representing actually occurs in all the Surface Contact Gestures because they imitate the body doing an action, in which the articulator stands for itself. The examples above were analysed with a specific focus on how an |object|’ shape is inferred by certain hand shapes or movement patterns. If the context was such that the interest was not on |object| being described but on a body movement being demonstrated, the conceptualisation would have been different. It would have been reduced to the articulator unit representing a complex body motion. The articulator would represent itself. The body movement itself would have a form that is to be identified as similar to an action known to the addressee. Note that the movement would not
have to be exactly equal to the action that it represents. It would just be required that the gesture was similar enough to the action for an addressee to recognise it.

For example in Figure 8.39(a,b) above, the articulating hands represent hands performing the action of gliding over the surface of the Drop. In Figure 8.39(c-e), the right hand represents a hand holding the object and the left hand finger tips represent finger tips tracing up to the tip of the same object (held in vertical orientation). The hands themselves imitate the action of hands holding and gliding over the surface of the Drop. So if the focus was strictly on the hands, for example, if the Drop was a very valuable and fragile object and the gesturer showed an interlocutor, how it should be handled—with two hands, with steady movements, and with spread out fingers so as to not drop it—the focus would tend to be on the three-dimensional shape and movement of the articulators, making it an example of Representing in three dimensions.

All Surface Contact Gestures include two form relations: (1) the one discussed at length, the articulators that disclose shape aspects of an object by a combination of gesture form operations, and, additionally, (2) the articulators instantiating body parts that perform certain actions. These two alternative foci co-exist for many of the analysed gestures.\(^{59}\) Besides depicting an object, the hands (or whatever the profiled articulator unit) needs to be conceptualised as hand in object-interaction (Müller 2014:1696). The choice of the terms Articulator Profiling and Shape Profiling stresses that, while a specific shape is profiled, the complete articulator and the whole body is not ignored by the observer, but rather it remains the base in relation to which the profiled shape is to be understood. In this sense the term ‘profiling’ closely follows Langacker’s use of “profile” (e.g., Langacker 2008:66-7, see Section 5.2).

The moving or static body as the base of any of these interpretations is crucial. Not only is it true that the visible body is the origin of all form interpretation, but also it is necessary to infer parts of the articulators that we cannot see (because they are occluded). For example in Figure 8.33b, an observer has to be able to recognises the whole-body performance and know what kind of shape is enclosed within the hands—without having immediate visual access to the hands’ inner surfaces. It is enough to see some aspects of the embodied action: the contact

\(^{59}\) See, e.g., Figure 8.29b, Figure 8.30b, Figure 8.31b, Figure 8.32b, Figure 8.33b, Figure 8.34b, Figure 8.36b, Figure 8.38, Figure 8.39, Figure 8.41.
with the object can be inferred from the rich tactile and visual experience that we associate with the performed action, as analysed by Mittelberg (2010b; Mittelberg & Waugh 2009).

Depending on the context, the focus can be more on the body, more on an |object|, or on both. If the task is to obtain information about an object, as is the case in Study 2, the focus is on the |object|, if the task is to learn movement patterns (as, e.g., in teaching swimming, dancing or body building), the focus is on the articulators that are the |body parts| participating in the respective action, and if the task is e.g., to mould dough or clay with your hands, the focus is more or less evenly split between the two. Cues in spatial and non-spatial contexts may be used to determine the focus of any gesture (for speech cues, see Mittelberg (2013); Talmy (2012; forthcoming)).

Mandel's (1977:59) example of gesturing "I look at my hand"—by looking at one’s own hand—is an example of several articulators representing themselves in their own large articulator unit. Not only does the hand instantiate the hand being looked at, but also the head and the eyes instantiate themselves as looking in the direction of the hand. Here action and the imitation of action can be very similar. This is a showcase example of Representing in three dimensions.

I also count Clark's (2003:243 et seq.) examples as Representing in three dimensions. In a rich investigation of “placing” one’s whole body or specific objects, Clark analyses different goals aspired by presenting something to someone. A specific action can fulfil the manipulatory goal of an action while also the communicative function of a gesture. For example, the act of placing an object on a check-out counter presents the object as an item-to-be-bought while also performing the goal-directed action of making it available for processing.

8.6.2. An articulator representing something else in three dimensions

Representing in three dimensions occurs only a handful of times in Study 2—as a creative way to characterise an object. The one participant in the pilot of Study 2 that also used her hand to stand for the Oval in two dimensions (Figure 8.27b) uses different articulators to represent an object three-dimensionally. While the pictures of Figure 8.45 show various gesture types that are employed to draw attention to her head, the head itself is used to stand for a spherical shape in three-dimensional representation, “if one would make my head twice as big” (“würde

---

60 These gestures are not analysed here since they are of other gesture types.
man meinen Kopf doppelt so groß machen”; the large sphere she refers to is not part of the nine objects of the final study).

In another object description, she brings her right-hand fist to a configuration that (surprisingly) comes close to the Egg and says laughingly, while Drawing around the horizontal outline of the Egg-shape hand with her left index finger (Figure 8.46), “but just with a constant form” (“aber eben mit einer konstanten Form”), laughingly. The occurrence of these creative gesture performances for rather simple forms might be due to the design of Study 2, in which the participants were not allowed to compare the stimulus object to other every-day objects such as eggs (see study design in Section 9.2-8).
In sign language, “partitioned” body parts (Dudis 2004:223 et seqq.) also act as surrogates (Liddell 2003:141 et seqq.) representing a variety of things as for example acting body parts, a vehicle, or a tree including branches (Figure 8.47, reproduced from Schembri et al. 2005).\textsuperscript{64} The articulator can be more or less iconic (i.e., similar) to the referent in its three-dimensional shape. While the cited authors focus on integrating a “real space model with […] the mental representation” to a “real space blend” (Liddell 2003:148) in sign language use, the thesis at hand specifically focusses on the form relations between the two. Besides this difference in focus, this approach resonates with Dudis’ and Liddell’s work (see also Part I, especially Chapter 6).

Figure 8.47 shows that, whether the iconicity is stronger (index-middle-finger legs), or weaker (index-middle-finger-thumb vehicle), the reference object is represented in all three dimensions (though in reduced scale), affording to depict movements and rotations in three dimensional space. Although for example the left side of a vehicle does not look very much like the left side of the hand in Figure 8.47, the left side of the hand will always—in the depiction of the vehicle—stand for the left side of the referred vehicle, defining it continuously, together with the rest of the hand, in three-dimensional rotation.

\begin{center}
\begin{tabular}{|c|c|}
\hline
Referent & ASL. \\
class & target \\
\hline
Animate & 2legs \\
Vehicle & 3edge \\
Airplane & ILY \\
\hline
\end{tabular}
\end{center}

\textsuperscript{64} The articulator head is also used to represent itself in showing facial expressions, see Dudis 2004.
8.6.3. Emblems

Emblematic gestures, short ‘emblems’, are often not entirely only arbitrary and not only conventionalised. For example, the German gesture of pointing at one’s own forehead with the index finger is (mildly) insulting, characterising the interlocutor (or a concept currently shared by both interlocutors) as stupid. This gesture is called “jemanden den/einen Vogel zeigen” (literally: ‘to show someone the/a bird’, Duden⁶⁴), which relates to the metaphor of a stupid person having a bird inside of the head (among other interpretations). In this understanding, it is, in part, a regular deictic gesture, Pointing At Object(contact).

However, also the arbitrary content of a gesture usually referred to as an emblem can also be categorised by Gesture Form Analysis. An emblem has to be perceived as similar enough to a specific conventionalised configuration and movement (‘how the gesture is performed correctly’) in order to be understood. If the ‘bird’ gesture is for example performed at the chest instead of at the forehead, would not be understood.⁶⁵ Thus, there is a specific configuration and movement pattern of the physical articulator associated particularly with an entirely arbitrary, conventionalised gesture. The gesture has to be articulated well enough in order for it to be communicative. A successful emblematic gesture includes mapping the physical, three-dimensional form of the articulators onto how the articulators ought to be moved in this gesture. This mapping between two three-dimensional forms is the reason for emblems belonging to the gesture type Representing in three dimensions. Since the correct form of an emblem does not specifically relate to a specific body (the gesturer’s or another person’s), this variant does not clearly belong to any of the two preceding headlines (Sections 8.5.1-2).

---

⁶⁴ [URL: www.duden.de/node/708844/revisions/1363839/view], last access 18.03.2016

⁶⁵ How much of a deviation (at the temple/at the cheek/at the ear) is still acceptable is an empirical question which is not part of this investigation.
8.7. Other gesture types: complex pointing and ‘minimal content gestures’

This section goes through some gesture types that have not yet been discussed, but are gesture form combinations that do occur—either in Study 1 or 2; or in other data analysed in gesture literature. They often do not follow the predominant schemes of either ‘line profile resulting in Pointing and Drawing, as discussed in Section 8.4 or ‘surface profile resulting in Surface Contact Gestures’, as discussed in Section 8.5. Instead, the gestures analysed in the following show how a surface profile can also be used to perform various types of pointing (Broad Pointing, Surface Directing, Section 8.7.1-2), and that also forms of more than zero dimensions can ‘draw’, more specifically imaginary leave traces whose shape is part of the communicated content (Line or a Surface Sketching, Sections 8.7.3-5). The chapter closes by analysing gestures whose form is only relevant to the overall utterance in a reduced, simplified way (Representing in oD, Section 8.7.6). If the following gesture types are performed similarly to the previously introduced types, cues for differentiation of their performance will be given.

Note the difference between alternative interpretations that lead to entirely different resulting forms and alternative interpretations that contain deviating constituent forms but have equal or similar resulting forms that fulfil the same function (e.g., Trace Leaving and Adjacency swapping places as in the torus example, Figure 8.41). The following sections will provide examples of both. Another discussion of different interpretations of the same articulator form can be found for example in Section 8.7.3.2. Another discussion of different interpretations of the same articulator form can be found for example in Section 8.7.3.2, but also in Wilkins’ treatment of “Etic Units” (Wilkins 2003:209).

8.7.1. Broad Pointing

Stereotypical pointing gestures profile the gesturing articulator as a line an continue this line until meeting the target. Yet some pointing do not involve a precise direction (one-dimensional shape profile), but point at a range of things or at a large area (sometimes approximately) either with a flat or an open hand. Kendon and Versante (2003) lay out analyses of index, thumb, and open-hand pointing, oriented palm up, down or oblique, concluding “that the form of pointing adopted by a speaker is systematically related to the way the object being referred to is presented in the speaker’s discourse” (p. 109). A discussion of possible connections of their analysis of articulator shape and discourse integration with a Gesture Form Analysis would be in place, but is not feasible in this dissertation. Also, Study 1 and 2 did not stimulate enough pointing gesture and not sufficiently diverse pointing gestures.
Apart from stereotypical one-dimensional pointing gestures, the profiled form in a pointing gesture can be two-dimensional but still predominantly extend in one direction. This form is extended away from the gesturer by the operation of Continuation, which entails that it also becomes wider the farther the shape is continued. In these cases, the resulting form is a plane in the form of a triangle with the narrowest angle directed towards the gesturer and both legs of the triangle delimiting the indicated area (Broad Directing (triangle), Figure 8.48a). At the far end, the triangle continues infinitely or is at least not clearly defined. In a related type, the triangular plane intersects with the surface of another object, for example a book shelf, in a line segment (Broad Pointing (triangle), Figure 8.48b). This line segment could, for example, highlight a horizontal segment within a long row of books, draw attention to one shelf in contrast to another higher or lower shelf, or single out one specific book, in case there is only one book per shelf. In general, the wider the flat hand is spread, the wider is the range that is indicated with these two gesture types. For Broad Pointing, both the spread of the hand and the distance of the intersection determine the size of the resulting line.

Both pointing types can be compared to water coming out of a squeezed water hose. When the tip of a hose is squeezed flat, water squirts out in a triangular plane (Figure 8.48a). When the hose further is directed against a wall, the water will hit the wall along a line (Figure 8.48b).

Another related type of pointing can be performed by pointing with the open hand, palm down, fingers loosely extended and directed forward, thumb not raise to height of the other fingers. This pointing gesture, Broad Directing (cone) results in a three-dimensional form
shaped like a cone (Figure 8.49a). Again, the resulting form of Broad Directing (cone) could intersect with a surface to mark an area on this surface, constituting the gesture Broad Pointing (cone). This area would be of a circular shape (Figure 8.49b).

An analogy for these types of pointing is liquid squirting out of a spray can. When sprayed in the air, the paint particles will travel through the air (during the first couple of centimetres) in the form of a cone (Figure 8.49a). When sprayed against a surface, the paint will cover a surface in the form of a circular blot (b).

In these gesture types, Adjacency can be last in the order of operations (as presented in Figure 8.49), but it can also move one position to the left (a) or up to two positions to the left (b), as indicated by the arrows. The intermediate constituent forms would change in shape and dimensionality, but the resulting form would remain the same. Gesture form interpretations of different orders among the last one or two operations (as indicated by the arrows) will be counted as a single gesture type, so that the number of gesture types discussed so far is four.

In the analyses of Wilkins (2003) and Streeck (2009), we find Broad Pointing (cone) directed at an area that begins already on the ground directly in front of the speaker—sometimes extending to areas far away from the speaker. Figure 8.50 shows this application of Broad

Figure 8.49: (a) Broad Directing (cone) and (b) Broad Pointing (cone).

64 The form of a cone is not to be confused with the use of a cone to capture the degree of precision of a pointing gesture (Kranstedt et al. 2006) though both cones capture marking not a point, but a larger space. Kranstedt et al. use the geometric shape of a cone to account for imprecision within trying to point precisely. For example in Figure 8.48-Figure 8.52, the pointing is intended to mark a larger area.
Pointing (cone) schematically, while Figure 8.51 shows empirical data from Arrernte speakers in a study by Wilkins (2003; for another example, see also Streeck 2009:79, ‘Fig. 4.10.4’, right person).

![Figure 8.50: Broad Pointing (cone) directed towards the ground in front of the speaker.](image)

The gestures reproduced in Figure 8.51 can be interpreted as examples of Broad Pointing (cone). Since the hand in Figure 8.51c seems to rather be spread out flat on a horizontal plane, this gesture could alternatively be interpreted as Broad Directing (triangle), but this different conceptualisation would result in indicating roughly the same area of land. This similarity in function between Broad Directing (triangle) and Broad Pointing (cone) shows that some types that are structurally different create forms that are not markedly different (both Figure 8.50 and Figure 8.52 both indicate the ground in front of the gesturer). The closer the cone in Broad Pointing (cone) is to the ground and the wider the angle of the cone, the more similar is the resulting form to that of Broad Directing (triangle): the ellipse shown in Figure 8.50 will become larger and increasingly open up at the far end. When the pointing direction of Broad
Pointing (cone) is parallel to the ground, the resulting form can be even more similar to Broad Directing (triangle; Figure 8.48a). Further, if we conceptualise space as an area (2D), as it is often done in talking about areas of land, the distinctions of the triangular and conic types of broad pointing vanishes altogether (both resulting in what is schematically illustrated in Figure 8.52). The kinds of pointing as shown in Figure 8.48 to Figure 8.52 are referred to as Broad Pointing, a group of gesture types that are not necessarily exactly equal in form, but similar enough to serve the function of marking space in a broad, sometimes vague fashion. Groups of types that exhibit only small differences in the resulting form and who essentially have the same function are called ‘functionally homogeneous types’.

The specific movement execution of Broad Pointing (cone), see Figure 8.49 through Figure 8.51, may include lateral or circular motion. This motion could be understood as leaving a trace, thus broadening the highlighted space according to the operation of Trace Leaving (Section 7.2.9.). Regarding the order of operations, Trace Leaving is similar to Adjacency in its variable position in the sequence of spatial processing analysed in Gesture Form Analysis. It can be inserted after one of the last four operations (including Adjacency).

---

65 In Broad Directing (triangle), this form is the output of the operation Continuation, whereas in Broad Pointing (cone), it is the output of Adjacency (Enclosure) or Intersection (see Figure 8.48 and Figure 8.49).

66 E.g., Maps are highly two-dimensional. We speak of ‘areas’ of space although they obviously consist of 3D materials and objects; we measure land, apartments, or parking space in square (kilo)meters, etc.
Some Aspects of Wilkins’ analysis bear resemblance with some gesture form operations. In particular, the effect of motion on gesture conceptualisation is described in a way that is highly consonant with the spatial relation captured by the operation of Trace Leaving: “When identifying large regions, or indicating the degree of spreadness of a mass object (like water), this point often co-occurs with a sweeping, rotating motion” (Wilkins 2003:195).

8.7.2. Surface Directing

Surface Directing is not part of the group of Broad Pointing but a pointing type functionally homogeneous with Directing, at least for gesturers standing on even ground. Surface Directing is equivalent to Broad Pointing in that the hand is held flat, fingers spread or held together, but this equivalence pertains to the level of a form parameter only (configuration), which does not mean that this gesture, and its constituent forms—specifically the resulting form—is conceptualised in the same way. A Gesture Form Analysis is proposed to disentangle structural differences and similarities to other pointing types. It will be shown that the structure of gesture form in Surface Directing is different to that in Broad Pointing and so are the functions of Surface Directing—despite of the similar hand configuration. This account parallels the system of pointing gestures in Arrernte speakers, among whom a difference in pointing functions is manifested by different words in their language. Wilkins (2003) identifies “Wide Hand Point” (p. 194) that widely overlaps with Broad Pointing contrasting the “Flat Hand Point” (p. 196), that overlaps in part with Surface Directing.

Surface Directing unites a surface profile (flat hand) with a precise indication of a direction. This could be a controversy because a surface extends in two orthogonal directions. Even if there is one dominant axis, the degree to which the articulator also extends on a second axis should be the degree to which this gesture is imprecise in giving a direction, as just discussed for Broad Pointing. But this is not the case.

Surface Directing makes use of another important form that the context of human gesturing often provides (see Gibson’s affordances of the ground, Section 7.1.2), namely the surface of the earth, and conveys information about a direction to walk or drive specifically on this surface. The Intersection of the surface that is continued from the profiled surface and the surface of the ground is a line: the precise indication of direction. Alternatively, the linear form as the result of the intersection can also refer to the predominant extension of an oblong referent. This is the case in the first of two examples by Wilkins in Figure 8.53a, where the orientation and possibly also the location of a mountain range is indicated.
When indicating directions and giving route descriptions, Surface Directing is similar to Directing (see gesture indicating North, Figure 8.10b) and almost fulfills the same array of functions. Figure 8.53b shows an Arrernte gesturer indicating South by Surface Directing.

Wilkins’ documentation of the example reproduced in Figure 8.53a stands out in that it includes a partial view of the direction pointed into. More commonly in gesture literature, the camera is directed frontally at the participant, giving little impression of the area the pointing gesture is directed toward. Since the mountain range is not discernible on the image, Figure 8.54 is offered as a schematic illustration of Surface Directing from a side perspective that allows seeing the resulting form in relation to the ground. It also gives an impression of why a surface profile has advantages over single-finger Directing: when the path to be taken is not conceptualised as completely level, the intersection follows every dent, bump, hill or valley.
Thus, Surface Directing creates a form that strictly extends into one direction on the horizontal plane, but it creates vertically flexible path. It follows that this gesture does not allow indicating for example the desired flight direction of an airplane including the information on ascent or descent.

The vertical extent of the articulator not being a by-product but a constitutive part of Surface Directing is underscored by the motion that is often part of this gesture. As shown in Figure 8.55a, this gesture is often performed with a vertical movement, in this case an arced motion, to the front as well as up and down. The vertical motion emphasises the vertical extent of the flat hand shape. Since the output form of the Trace Leaving operation remains of the same dimension—given parallel, co-planar, motion—the form remains two-dimensional while making the vertical extent of this form more salient (by redundancy with the vertical extent of the hand shape itself).

Surface directing occurred often in a warming-up question of Study 2. The participants were asked to explain how they got from the main entrance of the building to the Natural Media Lab on the fourth floor, a demanding endeavour for people not yet familiar with the building. Figure 8.55b shows Surface Directing applied for the indication of a path that includes going up: walking up a staircase whose stairs are ascending towards the right (“then we went right, up the stairs”/”dann sind wir rechts die Treppe hoch”).
The flat hand is profiled as a surface (green in Figure 8.55b). By moving the hand forward and up and down in parallel to the surface of the hand, this surface is enlarged, as it leaves a trace spanning a larger area (in pink) than what could be represented statically by one hand. This larger surface, extending predominantly to the front is continued away from the gesturer (orange). The intersection of this surface with the ground of the imaginary scene is the path information that is the resulting form of this example of Surface Directing. The room this gesture refers to (the lobby of the building) contains a staircase on the far right, as mentioned in the accompanying speech. The intersection path of the continued surface (produced by the gesture) and the ground of the |lobby| adapts to elevations of the ground such as the stairs of a staircase, as illustrated in Figure 8.55b.

Surface Directing can also be used to depict a spatial arrangement that is to be understood as a metaphor. The work of Müller (2008b:228-33) features an example that could be categorised as belonging to the gesture type Surface Directing. A participant argues that “the first job after finishing the education does set the tracks for future jobs” (Müller 2008b:229). He gesturally indicates different directions in space, which represent different career paths (of whether to study medicine, German studies, agriculture or become a tennis teacher). The gestures are performed differently, but share a vertically orientated articulator (as in the
preceding examples Figure 8.53–Figure 8.55), be it a single hand palm vertical, two vertical hands behind each other roughly on the same plane, or two vertical hands in palm contact. The gestures are accompanied by the German expression of ‘setting tracks’ ("Weichen stellen"), meaning to make a crucial decision for one’s future career. The concept of setting tracks does not target a specific goal (where a Pointing At gesture could be expected) but indicates the different directions along which one can travel. In analogy, job choices are conceptualised as career options leading to different life realities.

The screenshots in Figure 8.56 reproduced from Wilkins (2003) and Fricke (2007) exemplify Surface Directing of a curved or bent path. Figure 8.56a displays an Arrernte speaker giving directions in a car slightly flapping down the fingers of her left hand; Figure 8.56 shows a German participant performing a “direction-deictic gesture” (Fricke 2002:209) as part of a route description, bending the palm and fingers towards her left. To allow for a clear indication of a path (as schematically illustrated in Figure 8.54), despite the exhibited curvature, the profiled surface should be curved on its horizontal axis (but not in its vertical axis). The examples in Figure 8.56 meet this prediction. The figure shows the hands curved horizontally in a way that stands for the degree of curvature of the path to be taken, but not curved vertically. Vertical curvature would make the profiled surface oblique in relation to the ground, thus hampering a clearly defined intersection (which is a central feature of the gesture type Surface Directing).

Figure 8.56: Examples of Surface Directing in route descriptions (a) from Wilkins (2003:197; excerpt) and (b) from Fricke (2007:110; 2013:741, 2014:1809; excerpt). (a) Courtesy of David Wilkins and Taylor and Francis Group. Copyright © 2003, Taylor and Francis Group LLC Books. (b) Courtesy of Ellen Fricke (www.ellenfricke.de) and De Gruyter. Copyright © 2003, Walter De Gruyter GmbH. Through Copyright Clearance Center.
There is quantitative evidence for a co-occurrence of gestures that fulfil the function of indicating direction and the hand shape of a vertical flat hand, palm lateral (see also Haviland 2003:160 for Tzotzil speakers): Fricke, identifying gestures as shown in Figure 8.56b as ‘direction-deictic’ (“richtungsdeiktische Geste”, Fricke 2007:110) found 80% of the gestures accompanying the term “straight ahead” (“geradeaus”) in a route description task to be performed with an open hand palm lateral (Fricke 2007: 111, see also 2013:741). Of 16 Arrernte speakers, 100% used the flat hand vertical to point North (Wilkins 2003:196).

In the next sections, surfaces will be used to create a different kind of resulting form: not a lower-dimensional form (1D) as in Surface Directing, but a higher-dimensional form (3D) that can be used to point in a broad fashion or sketch a volume of space.

8.7.3 Line Sketching and Surface Sketching

In order to analyse Line Sketching and Surface Sketching, some differentiations within the group of gestures that can be called sketching and the gesture form operation Trace Leaving are necessary. Trace Leaving is the driving operation in the gestural strategy Mandel calls “sketching” (1977). In this regard, the gesture type Drawing actually also belongs into the category of sketching, since Drawing includes extending the dimension of a constituent form, the point, to a line by interpreting it as leaving a trace (e.g., Figure 8.22b). Under the name of point-sketching, Drawing fits perfectly as the simplest kind of sketching (since the input form for the operation of Trace leaving is a zero-dimensional point). But Drawing is also closely related to pointing, overlapping in up to three operations: (1) the profiled shape is one-dimensional, (2) it can be continued to point/draw in the distance, and (3) it intersects with a surface producing a point as output form. The membership of Drawing in the group of Pointing as well as in the group of sketching is an example for problems of categorisation that cannot be solved at the level of exclusive categories but only at the level of a matrix of gesture form operations. Drawing simply involves operations that are salient in pointing gestures as well as operations that are salient in sketching gestures. This follows Kendon’s insight that gestures can also occur in which that feature of it that is seen as “pointing” is combined with other features, whether of hand shape or of movement, so that the gesture not only indicates but at the same time accomplishes something else, such as depicting a characteristic of the object indicated.

Kendon 2003:312
Other gesture types that fall into Mandel’s sketching category were also already analysed, for example Moulding. In Moulding, the two-dimensional surface leaves a two-dimensional trace, since it is moved in parallel (co-planar) fashion along the extension of the profiled surface. If the profiled surface is plane, the movement can infinitely extend as long as it remains on this plane. If it is curved, it can be moved following the given curvature.

This kind of movement along the extension of the profiled shape sometimes also occurs in simple index-finger pointing gestures, for example the finger profiled as a line is being moved forward, parallel (co-linear) to its own extension, at the end of a pointing movement. Kendon & Versante state about the movements of a pointing gesture “that at least the final path of the movement is linear” (2003:111). This parallel movement emphasises the direction pointed into, it does not change the indicated direction or location. The use of form redundancy within the constituent forms of a gesture as a means of emphasis and reaffirmation is, I think, a pervasive strategy in gesture conceptualisation that occurs in many gesture types.

In contrast to the mentioned cases of parallel (co-linear and co-planar) applications of Trace Leaving, the same operation produces different forms when there is a lateral movement in relation to the profiled shape (e.g., orthogonally or obliquely). The second row of pictures in Figure 8.57 shows the output of a point, a line and a plane being moved laterally (orthogonally or obliquely), each leaving a trace of exactly one additional dimension.

Figure 8.57: Applying the gesture form operation of Trace Leaving with a movement orthogonal to the profiled shape, resulting in different types of sketching, including Drawing.
The following sections will discuss the gesture types illustrated in Figure 8.57, proposing an explanation why, except for Drawing, sketching gestures may occur less than the alternative types involving Trace Leaving, above (e.g., Moulding).

8.7.3.1. Line Sketching

There are multiple speculative explanations why Study 2 did not contain clear-cut examples of Line Sketching (sketching a surface with a line). First, there is an intrinsic problem of this gesture type that favours competing types like Moulding (which can also create a surface): As shown in Figure 8.57, Line Sketching cannot easily be distinguished from the more commonly used Drawing (Point Sketching) because the hand shape is the same. This is not only true when using the index finger for the sketching, but actually when using any articulator, since gesture form is articulator-independent (it can even be an oblong object, e.g., a pointing stick). Thus, using Line Sketching as a gestural depiction will always be in danger to be misinterpreted as its more prominent relative Drawing, since configuration and movement can be the same.

Another reason for the scarcity of Line Sketching is the fact that gestures creating a surface as constituent or resulting form can more easily be performed with an open hand, a surface-like articulator (e.g., Figure 8.39). In using the open hand, the articulator already stands for the form of a surface by itself. Moving the articulator parallelly (co-planar) to its profiled surface additionally emphasises and elaborates the depicted surface. Mutually affirming, both the curvature of the hand and the trajectory of the movement depict a surface (in Line Sketching, the movement alone would have to convey the surface). Summing up, there is no need to run the risk of confusion with a Drawing gesture, if the open hand serves this purpose more readily.

Nonetheless, Line Sketching does occur, apparently as a rarely used type. In the data of the dissertation study it only occurs in one gesture of a participant when she is holding the two index fingers at some distance, pointing at each other, and moving them in parallel in a circle to depict a cylinder. This exception further supports the rare occurrence of Line Sketching: the two fingers are directed towards one another specifically to support imagining a line between the fingertips that is being moved laterally. Seemingly, only in a very special configuration that supports imagining a line between the two index fingers, the interpretation of a movement as Line Sketching may be favoured.
8.7.3.2. **Surface Sketching**

The flat hand “can leave a solid trace perpendicular to itself, as if one indicated the height of a stack of papers by putting his hand palm down on the table and then raising it to the height of the stack” (Mandel 1977:67). Also *vice versa*, a horizontal movement of a flat hand held vertically can sweep “out laterally to indicate a sector or quadrant [...] it can be used to delimit an area in which a search for food should take place” (Wilkins 2003:196-8). The cited kinds of gestures show the possibility to Surface Sketch vertically (Mandel) and horizontally (Wilkins), while Mandel (1977:67) identifies a constraint of Surface Sketching in *relative* orientation: the direction of movement has to be perpendicular—or, it may be added, at least non-parallel (non-planar)—to the extension of the articulator.

Surface Sketching comprises various interpretation variants that serve a very similar function. They all involve the articulator being profiled as a surface and the operation of Trace Leaving. Within the constituent forms of Surface Sketching, we find examples of a surface leaving a volume as its trace (Figure 8.58a,c) and an intersecting line leaving a surface as its trace (b,d).\(^{67}\)

Figure 8.58 documents horizontal (lateral) Surface Sketching. A city guide on a Copenhagen boat tour explains the reason for the typical proportions of houses: “this is also why we s- we have [intelligible] see so many narrow but very tall buildings”\(^{68}\). Located in front of the speaker are the seated tourists, so that the gesture does not directly point to physically present houses, but rather creates imaginary house miniatures in the guide’s gesture space. The analysed gesture occurs during the first words of the utterance, sketching a |row of houses| on which ensuing gestures will further elaborate by specifying the narrowness and height of single |houses|. Images (a-d) show several possible interpretations of the first single gesture which are presented in the following.

In the reading of Figure 8.58a, the sagittal-vertical extent of the flat left hand is profiled. Since the camera looks onto the back of the hand, the green profile is not visible. The surface is moved through space to the participant’s left, a motion roughly perpendicular to the profiled surface. Regarding the space travelled as the extension of the space marked in this gesture

\(^{67}\) This exemplifies that sketching a surface with a line may not generally be a rare conceptualisation in gesture, but that it does not occur very often on its own (Line Sketching) as argued in the preceding section.

\(^{68}\) The reason is that houses used to be taxed on the basis of the square meters they occupied, so that building narrow and high houses was cost efficient.
(Trace Leaving), the trace of the plane-shaped implement is a three-dimensional volume. Marking this volume in front of the gesturer as the solid |row of houses| is a possible conceptualisation of this gesture.

As an alternative to this interpretation, the |row of houses| could also be depicted by the (depth-less) |façade of the houses| (Figure 8.58b). This form would be the result of intersecting the plane of the hand with the frontal plane in front of the gestures, potentially at the finger tips. The existence of such an imaginary a frontal plane is consonant with Fricke’s ‘screen-like models’ ("bildschirmähnliche Modelle", Fricke 2007:267-8).

However, the precise location of the imaginary row of houses in gesture space is actually not conveyed in this gesture. The |houses| could also be assumed farther away from the gesturer (since one might ideally want be located at a considerable distance to capture the proportions of a big object such as a row of houses). The distance of either the imaginary solid |houses| (c) or the just the |façade| (d) would necessitate the additional operation of Continuation as illustrated in Figure 8.58c,d respectively. In fact, any other distance of the |houses| (or the |façade|) would be a possible conceptualisation, the |houses| could already begin at the wrist while also extending far into the depth of gesture space, or the profiled plane could be continued far away from the speaker to a distance (even much farther than in Figure 8.58d) as if the gesturer was looking from across the road or canal over to the houses.
While there is this multitude of interpretations, the main spatial concept of a row of buildings is captured in all these interpretations; the interpretations in Figure 8.58a-d are functionally homogeneous: The similarity in height of the |houses| is a result of the equal spread of the fingers throughout the movement (that leaves a trace). The arrangement of the |houses| is defined by the single direction of motion and of a single stroke (that leaves a trace). The vertical orientation of the |houses| is cued by the flat vertical hands (whose flat shape is profiled and continued). All the aspects of articulator form translate into the characteristics of the |row of houses| by the combinations of operations displayed in Figure 8.58a-d.

Note that, if the gesturer understood the gesture as illustrated in Figure 8.58a and her interlocutor conceptualised the same gesture as shown in Figure 8.58d, communication would most likely still be flawless. This kind of ambiguity in spatial conceptualisation seems to be a common product of gestural expression. A gestural depiction specifies only spatial characteristics of current interest, not all characteristics. In other words, gestures disambiguate forms that make a difference for the communicative act, and tend to not specify those forms that are irrelevant to the overall utterance. The contribution of Gesture Form Analysis in this dissertation is limited to laying out the possible interpretations of a gesture, and giving a structural description of the conceptual steps that make these interpretations possible.

69 Except a slight abduction of the thumb
8.7.4. Representing in zero dimensions

This and the next section treat gestures whose form is much less defined than in prior examples. I would like characterise these gestures by calling them ‘minimal content gestures’. For these gestures, the form is undefined to an extent that does not allow identifying complex derivations of articulator form. Instead, the bodily action can only be made sense of in a simplistic, even more abstract way. The articulator’s position and motion are abstracted away from the articulator (similar to “abstract motion” in Streeck, 2008:295). These gesture types are necessary because, in coding actual data, the gesturer’s articulators cannot be said to always represent a specific shape. A gesture might seem to be motivated but not executed in a manner that exhibits precise gesture form with regard to how the articulator is configured. With these much reduced examples of gesture form, this and the following section complete the whole spectrum of gesture form as investigated in this dissertation. While the following section (8.7.7) deals with the most extreme case of abstraction, the current section discusses one of the examples, in which only the shape of the articulator does not seem to be clearly defined while the position of the articulator or the change of position (motion) is perceived as meaningful.

Without motion

If only the position of the articulator is relevant, the articulator is conceptualised as a point, in other words a form lacking any shape specification. Still, its (remaining) position can be crucial information. It can stand for the position of a physical object and the positions of multiple articulators (e.g., both hand) can depict the arrangement of multiple objects. As in all prior gestures, this gesture type also includes the possibility of visualising of abstract entities.

Just raising the hand to a certain point in space—in a relaxed, non-marked laxly open hand shape—while introducing a new topic of conversation can be enough to be able to refer back to this topic just by bringing the hand to this position for a second time, or by pointing to this position. The mere position, a zero-dimensional form (standing for the new topic), is the spatial information conveyed by the first gesture that Represents in zero dimensions, and the spatial information reiterated by the second gesture Representing in zero dimensions.

With motion

An articulator conceptualised as a point can also indicate simple translational (not rotational) motion. Interpreting this motion as actual motion is possible. Alternatively, motion can be understood as leaving a one-dimensional trace, as in Drawing (but without an oblong
articulator; *e.g.*, a lax open hand instead of a stretched out index finger). In summary, this type is tightly bound to the affordances of a point (Section 7.1.2).

A common example is ‘pointing’ by moving an articulator in a certain direction. For example the gestural expression of “distant past, ‘Years ago!’” can be “evoked by the hand and/or head raised high backward, referring to a place far behind oneself” (Calbris 1990:120). The head is to be profiled as a zero-dimensional shape and its movement backwards in space metaphorically refers to a point in time long past; either simply by this point moving backwards, or by this point leaving a trace in form of a one-dimensional line that ‘points’ backwards. This gesture is to be observed in different cultural communities (Calbris 1990:120; Olanike 2009:254).

### 8.7.5 Indefinable Activity

Indefinable Activity resembles Representing in zero dimensions in that the shape of the articulator is not relevant. Additionally, the motion trajectory is not relevant either. It can thus be doubted that this is still inside of the domain of gesture form. I am including it for the sake of completing the targeted typology (Figure 8.59). Observers tend to identify a movement as motivated even though they can not specify any spatial information that is used to convey meaning. Beat gestures that only contain “rapid flicks of the fingers or hand” (McNeill 1992:80) may be examples of this gesture type. Still, an observer can see the gesturer’s effort of expressing something and the fact that a gesture is performed at a certain moment marks this specific point in time. The coinciding syllable, word or phrase can be highlighted. Activity during a pause in speech can help to maintain the floor and an active gesturer in general can be presumed to be more engaged than an inactive one. For these purposes, a specific gesture form is not necessary. For an occurrence of Indefinable Activity it suffices for an observer to identify an intentional, gestural movement. In other words, the gesturer has to be recognised as giving information in contrast to “giving off” information (Goffman 1963:13-4; Kendon 2004:7).

Many non-pronounced movements could serve as examples for this category, even very small movements. For example the first participant of Study 1 usually rested his arms and hands on his thighs when sitting on the stool in the recording space. In his first utterances expressing uncertainty about how to describe the first object, he repeatedly lifted only the thumb of the left hand up for a brief moment. While this movement does not convey specific spatial information, it is an activity that can be interpreted as meaningful. It could stress what
is being said or convey information about the truncated effort to perform a gesture—potentially as a means to show oneself willing but unable to gesture (or speak) more precisely at this moment. Whatever the further interpretation, neither the position nor any aspect of the movement seem to carry spatial information. There is no gesture form, but, potentially meaningful, ‘indefinable activity’.

8.8. A typology of basic gestural semantics based on Gesture Form Analysis

Section 7.1.2 introduced dimensions of form that afford acting upon space, e.g., indicating a location or dividing portions of space. Rooted in these affordances, Figure 8.59 shows a typology of gestures sorted by the dimensions of the ‘profiled shapes’ (central column) excluding various gesture types’ possibility of conceptually reducing gesture space to one or two dimensions.

The left side of this figure contains the denomination of the gestures: the group of gestures, in broad functional classes (first column) and the technical terms of the gesture types (second column). Column three shows the different variants within one gesture type (left-aligned), different types of motion or other comments (right-aligned). The right side shows the modular structure of the typology, in which the operations are ‘switched on and off’, thus spanning all gesture types analysed here. The rightmost column summarises the structure of gesture form (as elaborated in Sections 8.4–7) while also specifying the order of operations.

Gesture Form Analysis allows categorising gestures in a modular typology, strictly on the basis of gesture form—gesture form consisting of a sequence of imaginary forms required for understanding a specific gesture. The offered typology attempts to categorise all gestures (all visual gesture perception of ‘rough movement’, Figure 1.1) and, with caveats regarding the order (Section 7.2.1), also gesture production. Gesture Form Analysis provides a tool for a systematic and operationalisable form analysis based on two principles: forms of different topological dimensions elicited by specific gesture form operations.
# Gesture typology (in 3D space)

**Gesture types** analysed in this dissertation, by profiled dimensions and modular **gesture form operations**

<table>
<thead>
<tr>
<th>Group</th>
<th>Gesture type</th>
<th>(specification)</th>
<th>Shape</th>
<th>Continuation</th>
<th>Intersection</th>
<th>Trace</th>
<th>Adjacency (all kinds)</th>
<th>Abbreviated Code</th>
</tr>
</thead>
<tbody>
<tr>
<td>A: Indefin. Activity</td>
<td>Indefinable Activity</td>
<td></td>
<td>none</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>IA</td>
</tr>
<tr>
<td>B: Representing</td>
<td>Representing in 0D</td>
<td>+/- motion</td>
<td>0D</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0D</td>
</tr>
<tr>
<td>B: Representing</td>
<td>Representing in 1D</td>
<td></td>
<td>1D</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1D</td>
</tr>
<tr>
<td>C: Pointing</td>
<td>Directing</td>
<td></td>
<td>1D Continuation</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1D+CN+IN</td>
</tr>
<tr>
<td></td>
<td>Pointing At Location (+/ contact)</td>
<td></td>
<td>1D Continuation</td>
<td>Intersection</td>
<td></td>
<td></td>
<td></td>
<td>1D+CN+IN+AD</td>
</tr>
<tr>
<td></td>
<td>Pointing At Object (+/ contact)</td>
<td></td>
<td>1D Continuation</td>
<td>Intersection</td>
<td></td>
<td></td>
<td></td>
<td>1D+CN+IN+TL</td>
</tr>
<tr>
<td>D: Sketching</td>
<td>Drawing (+/ contact)</td>
<td></td>
<td>1D Continuation</td>
<td>Intersection</td>
<td>Trace L.</td>
<td></td>
<td></td>
<td>1D+CN+IN+TL</td>
</tr>
<tr>
<td></td>
<td>Line Sketching (+/ contact) non-parallel motion</td>
<td></td>
<td>1D Continuation</td>
<td>Intersection</td>
<td>Trace L.</td>
<td></td>
<td></td>
<td>1D+CN+TL</td>
</tr>
<tr>
<td>B: Representing</td>
<td>Representing in 2D</td>
<td></td>
<td>2D</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>2D</td>
</tr>
<tr>
<td>C: Pointing</td>
<td>Surface Directing (triangle/cone)</td>
<td></td>
<td>2D Continuation</td>
<td>Intersection</td>
<td></td>
<td></td>
<td></td>
<td>2D+CN+IN</td>
</tr>
<tr>
<td></td>
<td>Broad Directing (triangle/cone)</td>
<td></td>
<td>2D Continuation</td>
<td>Intersection</td>
<td></td>
<td></td>
<td></td>
<td>2D+CN+AD</td>
</tr>
<tr>
<td>E: Surface Contact</td>
<td>Holding +/- motion</td>
<td></td>
<td>2D Continuation</td>
<td></td>
<td></td>
<td></td>
<td>Adjacency</td>
<td>2D+CN+AD</td>
</tr>
<tr>
<td>Gestures</td>
<td>Moulding parallel motion (p.m.)</td>
<td></td>
<td>2D Continuation</td>
<td>Trace L.</td>
<td>Adjacency</td>
<td></td>
<td></td>
<td>2D+CN+TL+AD</td>
</tr>
<tr>
<td>D: Sketching</td>
<td>Surface Sketching (+/ contact, +/- depth) non-p.m.</td>
<td></td>
<td>2D Continuation</td>
<td>Intersection</td>
<td>Trace L.</td>
<td></td>
<td></td>
<td>2D+CN+IN+TL</td>
</tr>
<tr>
<td>B: Representing</td>
<td>Representing in 3D</td>
<td>itself/not itself, +/- motion</td>
<td>3D</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>3D</td>
</tr>
</tbody>
</table>

grey = not applicable for all types
Part III:

Quantified testing of gesture form:
Measure and Shape Gestures
Part I and II of this dissertation deal with a general model of gesture form, investigating dimensional distinctions and geometric operations. Gesture Form Analyses were conducted on a series of gestures, motivating the establishment of a gesture typology. Each gesture type has multiple tokens. A gesture with a particular articulator form creates diverse forms of specific magnitudes, curvatures, and so forth. Drawing gestures were shown to contain forms of specific dimensions determined by its operations, but the length and shape of the resulting line drawn in the air were not determined. The neutrality of shape and magnitude is a feature of topological dimensions (Talmy 2000:223-5) and is the reason for how a gesture typology of the suggested kind could cover many or all gesture tokens of the same gesture type.

Part III, the empirical investigation of this dissertation, deals with two gestures of a specific structure of gesture form in empirical data, captured by video and motion-capture cameras. It treats the forms of two particular variants of the gesture type Holding: the Measure Gesture and the Shape Gesture.

As shown in the previous chapter, Holding is most pervasive in gestures within a particular variety of contexts. It includes the “palm-up-open-hand” (Figure 8.29), McNeill’s “Sylvester and Tweety” gesture (Figure 2.7 and Figure 8.30), and the many other gestures that are conceptualised as an articulator unit being in static surface contact with an object. The Measure and Shape Gesture are two variants of Holding that have similar functions in object description and are performed roughly with the same articulator unit. Sowa and colleagues claim that they have one and the same articulator form, “a C-shape hand” (Sowa & Wachsmuth 2005:146).

The aim of Part III of the dissertation is to pinpoint differences between the resulting forms of these two gestures. This will shed light onto manifold ways of partial Enclosure. Depending on the kind of Enclosure, the object will be specified only with respect to aspects of its size (e.g., height, diameter) or with respect to aspects of its size as well as its shape (e.g., round, right-angle corners).

While the previous parts of this dissertation dealt with structural features that do not necessarily require a test for perceived differences—no one would claim that, for example, Drawing is the same as Holding—Part III will test gestures that have a similar articulator form. The investigation begins with testing whether participants reliably distinguish different conceptualisations as predicted by the proposed model of Gesture Form Analysis.
To test for specifics in gesture form as defined in this dissertation, a new methodology is needed. Chapter 9 will lay out the methodology for the three-dimensional data of Study 2 in the attempt to contribute to best practices in quantitative gesture studies. Chapter 10 will use the thus-processed data to test for a difference in gesture form of Measure and Shape Gestures, connecting to the theory of Gesture Form Analysis.

9. Design, setup and technical preparation of an object-description study (Study 2)

The data collection in Study 2 was designed to satisfy the demands of two very different kinds of scientific investigation. On the one hand, the video data is to be qualitatively analysed by a Gesture Form Analysis and quantitatively coded as Measure and Shape Gestures by raters. On the other hand, motion-capture data is used for quantitative testing that requires undertaking a series of post-processing steps to attain indicative variables.

While Part II laid the ground for the qualitative analysis by introducing a theory of Gesture Form Analysis and the accompanying qualitative methodology, Chapter 9 will report on methods of collecting and post-processing the data of Study 2, focussing on motion-capture data. The new methods cannot only be applied for Gesture Form Analysis or motion-capture-aided gesture analyses. The methods are applicable for different kinds of analysis of human movement data and even time-series data in general. Their contribution does not lie in technical sophistication (the computations are simple), but in practicality and abstraction of complex movement to indicative and intuitive variables of rich three-dimensional data. The following list summarises the methods and tools.

1) Two methods of attaining hand-shape information regardless of position, orientation and motion of the hand as a whole:
   a) A basic but extendable set of motion-capture reflectors that helps reduce the system load in a motion-capture recording while still affording various kinds of analysis: the Minimal Marker Set;
   b) A coordinate system capturing positions of the fingers in relation to the palm: Hand Shape Coordinates.

2) A script analysing motion-capture data in Java (MoJa) that computes a series of variables on hand movement including all Hand Shape Coordinates on all three axes.
3) A plugin for the annotation software, Elan, that automatically generates annotations including beginning and end times on the basis of motion-capture data. Relying on a threshold criterion (e.g., speed lower than 10cm/sec defining a low-movement phase within gestural action), the Elan Plugin for Automatic Annotation (EPAA) objectively determines occurrence and duration of annotations making raters’ subjective interpretations unnecessary for this part of the process.

4) A Data-Extractor script (DEx) extracts and averages values of variables, for example those created by MoJa, during specific annotations, for example those generated by EPAA.

The methods and scripts are only one step towards best practices. They make time-series data accessible to gesture analysis by straight-forward computations. A follow-up project could combine the MoJa, EPAA and DEx scripts into a unified or standalone tool to systematize the process, but this exceeds the scope of this dissertation.

9.1. Experiences from Study 1 shaping the design of Study 2

Study 2 is based on that part of Study 1, in which participants describe objects and are confronted with recorded object descriptions. The design of Study 1 was not sufficiently goal-directed in the sense that it did not yield well-packaged short recordings of rich gestural data. The following section will discuss the sources of this shortcoming.

9.1.1. Technical problems (system load)

In Study 1, a large number of optical markers caused the motion-capture system to not run smoothly. The system often crashed, marker recognition was prone to errors and some markers were recorded with large gaps. The huge amount of data produced by the x, y and z coordinates of each marker made it difficult to ensure errorless data. Markers on neighbouring fingers were sometimes confused by the system, which did not only demand more time-consuming post-processing, but also lead to possible data corruption if marker cross-overs were left undiscovered. It also became clear that many of the markers were unnecessary, would not be used in the later quantitative analysis, and their exclusion from the marker set would eventually favour the reliability of the data. Section 9.7 introduces the Minimal Marker Set, a set of only six markers per hand (rather than the original 26 on the hands and 14 on the rest of the body) which reduces the system’s computational load while still capturing all necessary configurational data.
Another problem concerned the video description task. Since Gesture Form Analysis does not only describe the necessary conceptual processes in the observers but also in the gesturers themselves, the design of Study 1, which included watching and interpreting pre-recordings of a single participant gesturing, was suboptimal. Although the selected stimulus gestures were vivid, creative and numerous, they would only allow for conclusions about gestures performed by that one individual.

Study 2 was designed to use the participants’ own gestures, which were recorded in a prior task during the same session. In this way, the data collected was not that of a third-party observer, but the gesturers’ own interpretation of their own gestures (with a certain time lag). The participants themselves spontaneously produced the co-speech gestures five to twenty minutes prior, were aware of the context including the object-description task, and had immediate personal experiences of their individual style of (bodily) expression.

The participants of Study 1 were asked to explain what information they obtained from a mute video of an object description. Gesture analyses of this data were difficult, since the participants’ interpretation of gestural action was not clearly restricted to one specific gesture. Without a one-to-one mapping of one part of the participant’s report to one observed gesture, no claim on the conceptualisation of one gesture type was possible. The observational data gathered in Study 1, though containing a lot of data points, was not controlled enough to support qualitative claims as those derived from a Gesture Form Analysis and did not allow for quantitative testing.

In the observer task of Study 2, the participant is shown only one gesture at a time and asked for the spatial information on the object specifically conveyed by this gesture (e.g., size, curvature). There is thus an assignment of every observation to one time-delineated gesture.

Another cause for the motion-capture system to overload was the length of the recording trials. The trial length is indirectly determined by the participants’ gesture rate, since, usually, the more a participant gestures, the less recording time is necessary for obtaining the same number of gestures.
In Study 1, despite the spatial task of object descriptions, some participants hardly gestured. Several trials did not contain gestures at all. To obtain enough gestures for statistical analysis, the trials had to be longer. Since the data load and system load is not lower when the participant does not gesture (the motion-capture system does not distinguish markers at rest or in motion), this study produced extensive motion-capture data that was difficult to handle and process and at the same time bearing only little relevant information.

9.1.5. Conclusion for a new study design

Study 1 produced several gestural types that could be broken down to gesture form operations in single qualitative Gesture Form Analyses. Yet for a comparative experiment on specific gestures, the variety of tasks was too large and too heterogeneous. The object description task was selected as the most promising design to import into Study 2 (in a revised version): participants would watch mute videos of their own object descriptions and would be asked to report what spatial information they obtained from each gesture.

The greatest challenge for Study 2, as it was also for Study 1, would be finding a compromise between a more controlled and a more naturalistic study design (Figure 9.1).

<table>
<thead>
<tr>
<th>Controlled</th>
<th>Naturalistic</th>
</tr>
</thead>
<tbody>
<tr>
<td>Participants should produce many gestures.</td>
<td>The atmosphere should allow for spontaneous, natural communication.</td>
</tr>
<tr>
<td>Participants not gesturing at all should be avoided.</td>
<td>The production of gestures should not be cued, à la: “Please produce a gesture to show the height of the object!”</td>
</tr>
<tr>
<td>Trials should not exceed 20 seconds.</td>
<td>Participants should be free in the order, quantity and style of gesturing.</td>
</tr>
<tr>
<td></td>
<td>The object descriptions should not be predefined but leave room for creativity.</td>
</tr>
</tbody>
</table>

Figure 9.1: Study 1 demonstrated the partially competing maxims of a controlled vs. a naturalistic study design.

Further, a list of criteria should be met to allow for quantitative testing:

1. The number of markers should be minimised to reduce system load.
2. Two specific gestures should be chosen for comparison. Quantifiable criteria should make these gestures stand out from the rest of the data, other gestures and moments periods gesticulation.
3. The hypothesis on the differences between the gestures to be compared should be quantifiable.
9.2. **Study 2 in a nutshell**

23 participants were handed nine objects, one at a time, in randomised order. After each object was taken away out of sight, the participant was asked to describe the shape of the object they just held in their hands to the interviewer, within 20 seconds. The descriptions were recorded by 14 motion-capture cameras and four digital video cameras (totalling 207 trials containing 471 relevant observations: 191 Measure Gestures, 280 Shape Gestures, see Chapter 10). Each gestural movement was rated by the participant (viewing a mute video of their description five to twenty minutes after the description task), and later by three expert raters. The appendices give information on participants’ background, motion-capture settings and hosting of the data. Following an open access policy, data and software were subsequently made available to the research community under the following link.

[https://clarin.phonetik.uni-muenchen.de/BASRepository/Public/Corpora/NM-MoCap-Corpus/NM-MoCap-Corpus.1.php](https://clarin.phonetik.uni-muenchen.de/BASRepository/Public/Corpora/NM-MoCap-Corpus/NM-MoCap-Corpus.1.php).

9.3. **Participants**

The participants (11 women, 12 men) were academics or students in Aachen aged 19 to 64 years ($\bar{x} = 34.5$; $SD = 13.3$; Figure 9.2) and fluent in German. While the pool was relatively heterogeneous in terms of academic background (see Appendix 3), the number is too low to be representative for people at the RWTH Aachen University, let alone Aacheners, Germans or humans. Conclusions on the basis of this pool of participants can thus only show structures that would have to be more broadly tested for any general claim.

![Figure 9.2: Age distribution boxplot.](image)
9.4. Lab setup

As shown in Figure 9.3, the Natural Media Lab was set up to capture one person (sitting on the stool close to the camera shooting the picture) by motion-capture cameras distributed around the volume in a setup similar to what is shown in Figure 9.3.

![Figure 9.3: Capturing volume @nmLab (Photo: Matthias Priesters), see also Figure 9.4 for setup.](image)

The participant faced the interviewer and was filmed, apart from the motion-capture cameras, by four video cameras aimed from different angles: three from the front, one from the back (to include the interviewer). Cameras DV1 and DV2 in Figure 9.4 are regular digital video cameras; cameras HS1 and HS2 are high-speed cameras recording at 100 frames per second. The participant was asked to sit on a stool (without armrests or other constraints for free upper-body movement) in order to be well visible from all sides and to ensure a stable position within the capturing volume. A screen in front of the participant showed the video stimuli.
The recordings of DV1, HS1, HS2 and all motion-capture cameras were conducted by an operator ("Op." in Figure 9.4) who cued the recordings for every trial. Each participant described one object per trial (for further information on trials and sequence of execution, see protocol in Appendix 4). Camera DV2 ran uninterruptedly throughout each entire session. The object stimuli were hidden in a box, out of the participant’s sight.

9.5. Technical setup

The recording was done by digital video cameras and optical motion-capture cameras. The post-processing setup can be found in Appendix 6. The hardware and software used was as follows:

- Video camera:
  - 1x Sony HDR-XR500 high-definition camcorder (running the whole time)

- Vicon MX series optical motion capture system:
  - MX Giganet, MX Ultranet control units
  - 10x Vicon T10 & 4x Vicon MX13+ series infrared cameras (set at 100 frames/sec)
  - Vicon Nexus v1.7.1 recording/processing software
  - 1x Sony DCR-PC9E digital video camcorder
  - 2x Basler piA640-210gc high-speed cameras (set at 100 frames/sec)
9.6. Minimal Marker Set

To obtain information on body movements, optical motion-capture systems employ small spheres covered in a highly reflective material, so-called markers, applied with surgical tape on the skin of the participant’s relevant articulator. Prior studies (e.g., Study 1) and test runs showed that using too many markers per hand caused the digital representations of three-dimensional markers to be difficult to distinguish in post-processing.

The preparation for Study 2 included developing a Minimal Marker Set that would allow the system to acquire all necessary hand shape information. To capture palm orientation, three markers are placed on the back of the hand at the thumb, index and ring finger (at the metacarpophalangeal joints, see triangle in Figure 9.5). Prior testing showed that these three positions move very little during changes in hand shape and that a marker at the pinkie was not acceptable: when the pinkie is moved back and forth, this part of the hand bends back and forth considerably. These three markers thus completely define the hand’s orientation since three different positions on a rigid object (not arranged in a line) are sufficient to determine its precise three-dimensional rotation. In other words, these markers operationalise the parameter hand orientation with the precision of the capturing system, which lies in the range of millimetres (according to manufacturer information).

One marker each is placed on the distal phalanx of the thumb, index and pinkie, precisely on the proximal part of the nail, representing the position of these fingers. The reasoning of usage of the Minimal Marker Set rests on the assumption that, in many gestures, as in the investigated Measure and Shape Gestures (Chapter 10), middle and ring finger tend to move with the index or pinkie. It also rests on the assumption that the angle of the intermediate joints of the thumb, index and pinkie are predictable by the distance between the marker at the metacarpophalangeal joint and the tip of the finger as well as the angle of these two
markers in relation to the triangular plane representing the palm orientation. For studies in which these assumptions do not hold, the marker set is readily extendable. Results of an extended marker set would still be comparable to the Minimal Marker Set, as long as the six positions (Figure 9.5, see also Hassemer et al. 2011) are included.

Figure 9.5: The six markers of the Minimal Marker Set.

9.7. Object and video stimuli

Two sets of stimuli were used in Study 2, object stimuli and video stimuli. First, nine objects were handed to the participant for inspection and then taken away for each of the participant’s descriptions. After a short break, the nine video stimuli of the participant’s nine descriptions were presented, in a newly randomised order, on a computer screen.

The objects were approximately hand-sized (2–18 cm in their longest diameter), of a rigid material and uniformly orange in colour to allow participants to focus on shape only (Figure 9.6; detailed information on stimulus objects in Appendix 5). Apart from their different sizes, their shape was characterised by different kinds of curvature: OvalPlate, RoundPlate, and HollowCylinder were cylindrically curved (bent on only one axis); and Egg, Drop, CutOffPlate, BigSphere, SmallSphere, and Torus were spherically curved (bent on two axes); Torus and HollowCylinder had a hole.
The second stimulus set is a series of video clips. Each video stimulus is the recording of the participant’s own description of each of the nine objects as performed in the earlier description task. The video stimuli were taken from recordings of camera DV1 and shown on a large screen in front and to the right of the participant (Figure 9.4). The videos were roughly 20 seconds long, and their projection was controlled by the interviewer via a laptop mirroring the same image. The videos were shown embedded in the annotation software Elan, to document the participant’s answers live in form of annotations that are assigned precisely to the sequences shown to the participant. Figure 9.7 is a screenshot of what was shown to the participant.
9.8. Tasks

This section will describe the tasks of Study 2 in more detail. The protocol in Appendix 4 is to accompany this section in showing the organisation of trials and tasks and specifying questions and recording cues.

The problems faced in Study 1 could not be entirely resolved in early unsuccessful attempts to run Study 2 due to participants' low gesture rate. One of two design changes that solved this problem was to introduce a minimal element of role-play. The participant is asked to imagine the interviewer to be an extra-terrestrial alien who does not know everyday objects of this world (e.g., egg, box) measurement units (e.g., centimetres, fingers) and is ignorant of geometrical descriptions (cube, round, parallel, height). Apart from these constraints, the alien interviewer understands and speaks German normally.

The second design change that produced satisfactory gesture rates was asking the participant to provide a response within 20 seconds. These introductorily mentioned features created a relaxed atmosphere by a somewhat comical background story (several participants laughed upon hearing these instructions) while also raising the (time) pressure in a playful way. With the alien introduction, there were no more trials without gestures. The participants performed much higher gesture rates than in the preceding designs, without being specifically prompted. Brief but rich object descriptions with engaged bodily expression were recorded, condensed to trials that mostly did not exceed the approximately 20 seconds. The tasks of Study 2, detailed in the following, are divided into two parts (that both follow up on tasks from Study 1).

9.8.1. First part: object description task

Each participant was asked to provide consent and completed the Edinburgh Handedness Inventory (Oldfield 1971). In the lab, warming-up tasks included asking the participants how they felt, whether they had a stressful day and requesting two route descriptions: one of how they found their way to the site of the Natural Media Lab; the other one from the entrance door of the building through multiple possible staircases and elevators, to a side wing of the building and up to the lab on the fourth floor.

The main task of the first part is the object-description task. The participants are informed about the details of the task and the 20-second time restriction. The peculiarity of the alien role-play is introduced. The interviewer hands the participant one random object to look at and handle and offered the possibility of asking questions. After retrieving the object and
hiding it in a box, the interviewer asks, “What did the object look like?” which is the cue for the participant’s 20-second description and for the motion-capture operator to start recording. A “Thank you!” stops the participant’s description and cues the operator to end the trial. The other eight objects follow randomly. If the participant ends after less than 20 seconds and some parts of the object are not yet described, the interviewer followed up on the details regarding their shape or size. In case the participant breaks the alien role-play rules, the interviewer asks back, for example: “Sorry, how long is two centimetres?” or “I don’t understand. What does an egg look like?”

After this task, the participant is handed a flat object and is asked to hold it up with both hands in a vertical orientation; this served for later system calibration and finger measurement. A break followed, giving the motion-capture operator time to prepare the stimulus videos for the second part of Study 2.

9.8.2. Second part: observer task

In the main task of the second part, the observer task, the participants were shown video sequences of their own gestures, one gesture at a time. The interviewer selected the first sequence of one gesture and asked the participant whether they thought what the hands did had anything to do with an object. If so, they were asked to provide information about what size and curvature information the gesture conveys. This question was accompanied by showing pictures illustrating three axes (height, width, depth) and three types of roundness (circular, cylindrical, spherical).

To not confuse participants with the subject of a topological perspective, types of curvature were illustrated as shown in Figure 9.8. Each column exemplifies one type of ‘roundness’. The interviewer emphasised that for example the category of circular (a) includes bent lines that do not close to a circle. Further, the categories span all continuous distortions, for example an ellipse counts as circular (a), a cone counts as cylindrical (b), and an ellipsoid as spherical (c). Multiple choices were possible and encouraged by the interviewer, as in “I see circular as well as cylindrical curvature in this gesture”.
Figure 9.8: Illustration offering the participant to choose from three kinds of roundness: (a) circular, (b) cylindrical, (c) spherical.

After collecting and documenting these responses, the interviewer showed the next sequence, asking whether what the hands do now still belongs to the last clip. If so, the two annotations were merged. If not, the questions on size and curvature were repeated regarding the following sequence. Throughout the interview, the participants were repeatedly encouraged to correct the selected sequences if they felt they should be longer, shorter, merged or split.

The interviewer clarified and repeated that "the question is specifically what is conveyed by the hands in the video clip" they had watched, not what the actual object looks like (just in case the participant was certain about the object that is being described); or what they thought the gesture was supposed to be showing, but failed to do so. Irregularly, the interviewer asked back whether the participant was sure of their choice of for example types of curvatures to support the participant in always considering all options.

As expounded in the evaluation of Study 1 (Section 9.1), presenting the participants with the gestures they performed in the first part of the study is important for the perspective of the research question, which underlies Part III of this dissertation. While the focus of Gesture Form Analysis is on the observer perspective, the constituent forms and operations are valid from the point of view of the gesturer, too (except the order of operations: Section 7.2.1). Similarly, the observer data attained in Study 2 is also gesturer data because the participant is the first-person observer and the gesturer of the presented gesture. Participants of Study 2 know all the objects; they just held them in their hands. They recall aspects of their description efforts and their bodies were very recently engaged in performing the gestures they are now to interpret. This kind of mirroring of their own gestures offers a closer view on the gesturer’s perspective than a third-person observer or rater could provide. Further, watching the video of
oneself was a stimulating task that yielded engaged responses about what the participant felt was not properly shown in a gesture and why. As a complement to the participant’s perspective, a follow-up study of completely naïve observers is suitable.

9.9. Data post-processing

The post-processing of motion-capture data was time-consuming, not specifically because the computational operations were complex, but because even for simple operations, as the ones laid out below, there is no software available that interfaces with the recording software VICON Nexus, or with the annotation software Elan. Section 10.2 will show the work-flow that we developed as an attempt to bridge this gap for the study of manual gesture.

9.9.1. Hand Shape Coordinates

As part of the Minimal Marker Set (Section 9.7), the three markers on the back of the hand (palm index, palm-thumb, and palm-pinkie) identify the position and rotation of the hand and will be used to anchor the Hand Shape Coordinates. Palm-index and palm-thumb define the first axis. The second axis is orthogonal to the first axis and intersects the plane defined by all palm markers. The third axis is already completely defined in its rotation by being orthogonal to the two already established axes (it is inverted for the left hand to be symmetric). These three axes form a coordinate system originating at the marker at the metacarpophalangeal joint of the index finger (Figure 9.9).

The coordinate system is dynamic in that it translates and rotates with the hand: raw motion-capture data outputs coordinates that change when the hand is moved. The coordinate system yielding Hand Shape Coordinates moves with the hand, showing the same \( x \), \( y \), and \( z \) coordinates for the index finger when this finger remains equally stretched out as for example in the canonical index-finger-pointing hand shape—regardless of where the hand is moved and where it is directed. The coordinate system is static within the frame of reference of the hand and dynamic with regard to absolute space.
Figure 9.9: Hand Shape Coordinates for the right hand, oriented according to the 3-finger-rule.

The orientation of the z-axis is mirrored for both hands so that, on each hand, the three-finger rule informs about the direction and the orientation of all three axes: the thumb points in the direction of positive x-values; the index indicates positive y-values, and middle finger positive z-values (Figure 9.9).

Hand Shape Coordinates will now be specifically used for operationalising data for the hypothesis of this study, but they are also offered as a general tool to measure hand-shape information. A dynamic coordinate system can also be applied to measure other body configurations, for example measuring body posture in relation to the waist (Grandhi *et al.* 2013). Quadrants or other shape delimiters inside the coordinate system can be used for hypothesis testing. Other applications include the position of head or arm (in relation to the chest), legs (in relation to the waist) or feet (in relation to the shin).

Prerequisite for the dynamic coordinate system is to capture three markers on a rigid body that serves as a frame of reference. It is applicable to any motion-capture system that attains absolute positions of markers.

---

70 Matthias Priesters convinced me of the three finger rule as a means of illustrating the dynamic coordinate system.

71 If the hypothesis was that an overstretched index finger occurred more often when a participant accompanied her pointing gesture with words such as "precisely", "exactly", negative z values for the marker at the tip of the index finger in the temporal vicinity of the selected words would test this hypothesis.
9.9.2. **Motion-capture Java script (MoJa)**

A script analysing motion-capture data in Java (MoJa)\(^\text{72}\) computes the variables that are relevant for hypothesis testing but cannot be directly copied out of raw motion-capture data. Hand Shape Coordinates are one group of examples. Hand Shape Coordinates are the basis for angle calculations (Section 10.5.3), which are also performed by MoJa. Other MoJa variables include the velocity and acceleration of markers, the distance between two markers or the measuring of the size of the hand in the calibration task.

9.9.3. **Elan Plugin for Automatic Annotation (EPAA)**

Variables attained in Study 2 are time-series data that can be used to demarcate sequences in the video according to certain constraints, for example simple thresholds. The Elan Plugin for Automatic Annotation\(^\text{73}\) selects these sequences, stores them, and displays them directly in a new tier in the graphical interface of the annotation software, so that the user can apply all Elan functions to these annotations.

Study 2 involves selecting low-movement phases that were defined by a velocity of less than ten centimetres per second. This process was automatised by EPAA. During the following steps of data selection, it was possible to modify or delete these annotations. EPAA can be used to automatically create annotations on the basis of time-series data from any recording device (motion capture, force plates, eye tracking, etc.). In gesture studies, a precise description of the movement pattern of a specific group of gestures can be supported EPAA because it can help to automate parts of the gesture analysis and introduce objective criteria to distinguish different gesture groups (thereby reducing subjective judgements).

9.9.4. **Data Extractor (DEx)**

The Data Extractor\(^\text{74}\) is a script that combines the information of the annotations (when a relevant sequence begins and ends) with the relevant variable (values of position, angle or distance of one or more markers). DEx creates the data matrix required for statistical testing. It registers every Elan annotation as one observation, whose values are the averages of

\(^{72}\) MoJa was programmed by Marlon Meuters, Matthias Priesters, and Svetoslav Ivanov Evtimov at RWTH Aachen University.

\(^{73}\) EPAA was programmed by Bela Brenger, Han Sloetjes, Matthias Priesters, and Zhi Lee at RWTH Aachen University.

\(^{74}\) DEx was programmed by Svetoslav Ivanov Evtimov at RWTH Aachen University.
designated variables at this time. For example, the angle of the pinkie might be 76.54 degrees in one observation since the data points of this variable fluctuate around this value during a 3.21 second-long hold of a selected gesture.

9.10. Summarising the characteristics of Study 2

Chapter 9 showed the design of Study 2 that elicits free gesturing of participants that view and report on their own gestures afterwards. It also presented scripts and software that post-process the recorded motion-capture data and automatically demarcate annotations that meet constraints for a quantitative analysis. The presented methodology and its constituent methods are still a loose set of tools that can be used for various purposes. The orchestrated use of these tools will be documented in the following chapter, which will also derive one research question that tests the central assumption of the theory presented in the prior chapters.
10. Are Measure and Shape Gestures in Study 2 perceived differently?

This chapter introduces the two groups of gestures that will be used for quantitative comparison: Measure and Shape Gestures. After defining the research question, Gesture Form Analyses of these gestures will show the differences in gesture form and predict differences in articulator form. These articulator form predictions will be operationalised to be testable in motion-capture data. The interpretations by participants and raters of all tokens of Measure and Shape Gestures will be attained, distinguishing both groups of gestures. After describing the steps of data selection, which draw on methods introduced in Section 9.9, the statistical results will be presented.

10.1. Motivation

It is known that the description of articulator form is a laborious endeavour both in gesture studies (e.g., Bressem 2013) and Sign Linguistics (Johnston and Schembri 2007; Hanke 2004). Since the concept of gesture form deals with the complex cognitive processes of a gesture observer, Chapter 7 only made claims about what is necessary in conceptual processing, rather than trying to account for the complete cognitive richness of a participant’s thoughts when observing for example Holding.

The empirical study of gesture from an observer perspective will always involve this conceptual richness. It is thus difficult to control for two conditions that only differ in exactly one aspect, while still allowing the participants to behave naturally and gesture in an unconstrained way. Study 2 tries to attain clear distinctions of what kind of spatial information observers associate with certain gestures. Statistical analysis poses another challenge: The operationalisation and quantification of both body movements (articulator form) and of gesture perception and interpretation (gesture form).

To face both challenges, Study 2 concentrates on two groups of gestures that (1) belong to the same type, (2) overlap in meaning, and (3) share similarities in articulator form. Nevertheless, according to Gesture Form Analysis, they differ significantly in how they are perceived; they are therefore different in gesture form. This difference will be made accessible by asking the participant questions that make a binary distinction between two specific gestures.

Study 2 complements the theoretical character of the presentation of principles of gesture form (Chapter 7) and the broad perspective of integrating all investigated gestures into a typology (Chapter 8). An empirical study should be more specific, also because the fact that
gesture types such as Drawing and Moulding are different is already generally acknowledged. As for the proposal of a basic theory of a Gesture Form Analysis, Study 2 cannot test the complex structure including all interactions between dimensions and operations of this model. It can only exemplify a distinction that is central to the theory:

In a nutshell, the line of argument begins with the general distinction between articulator form and gesture form. To derive meaning, observers conceptualise articulators as a specific kind of abstract form. This form gives rise to other forms according to the proposed operations, stepwise producing all constituent forms, as demonstrated in the conducted Gesture Form Analyses. The resulting form of these inferences contains the central spatial information, which is mapped onto concepts of physical or abstract entities. Taking two specific groups of gestures as an example, the distinction between articulator form (motion-capture data) and gesture form (as reported by the participants and independent raters) is tested in Study 2.

Since abstract concepts are more difficult to control for, the study is situated in the domain of object descriptions. In the two groups of gestures introduced in the following section, gesture form is to map onto physical forms, thus being easier to corroborate.

10.2. Similarity of Measure and Shape Gestures

Sowa and colleagues (Sowa 2006:146, 198-9; Sowa & Wachsmuth 2005: 146) regard two groups of gestures, in this dissertation called Measure and Shape Gestures, as equal in their articulator form (hand shapes reproduced in Figure 10.1). According to their analysis, both gestures are ‘‡”ˆ"‡†™‹–Š‹†‡šƒ†–Š—“އ†ƒ†ˆ"‹‰ƒǮǯǡ the other fingers spread.

Sowa does not only acknowledge these two gestures to be different with regard to gesture form. He also shows methods to compute referent shape from hand shape data (Figure 10.1, right side), for these and many other gestures. His concept of a “focus area” (grey in Figure 10.1) as the “meaning-bearing part of the hand” (Sowa 2006:91) seems to be equivalent to the surface profile in Holding gestures as analysed here. The only difference being that the operation of Shape Profiling in the framework at hand spans other form dimensions as well, for example linear profiles in pointing gestures. Sowa does not use the terms articulator or gesture form, but their implicit differentiation is part of his analysis and results. Including Sowa’s explicit calculations, his approach is very much in line with this dissertation (unfortunately, to my knowledge, his work has not been further developed).
While I agree that it is possible that one and the same articulator form can be interpreted as Measure as well as Shape Gesture, I predict the natural execution of these gestures to produce cues, in articulator form, that systematically distinguish the two. I will search for these cues by predicting differences based on Gesture Form Analysis and testing for these differences in the articulator form of specific gesture tokens—via motion-capture data. These cues will be specified in the hypothesis and operationalised (using the methods introduced above).

10.3. **Gesture Form Analysis of Measure and Shape Gestures**

Gesture Form Analyses of data from Study 1 drew attention to two related kinds of gestures that might lend support to basic assumptions of the theory of gesture form. Both kinds of gestures are independent-hold gestures; their expressive phase is a hold (Kita et al. 1998:27-8). In other words, these gestures are interpretable without motion, making them easier to illustrate by video-stills and graphical analyses (including the one that is part of Gesture Form Analysis) and easier to locate in three-dimensional motion data.

The gestures were chosen because of their similarity. Nevertheless they differ in conceptualisation, and Gesture Form Analysis helps to identify and operationalize this difference. If the participants systematically differentiated between these gestures, this could be taken as evidence that structural differences on this very fine-grained level are indeed being perceived by observers. In case the two groups of gestures showed specific differences in their articulator form that overlap with predictions of a Gesture Form Analysis, these specifications could be cues for one or the other gesture. If these cues could be broken down to specific
patterns in motion-capture data, this could even make these gestures candidates for objective (and automatable) coding criteria.

10.3.1. Measure Gesture

As discussed in Section 8.5.2, the gesture type Holding comprises gestures that imitate a hand being in surface contact with an object. The gesture in Figure 10.2 belongs to one group within this type, which will be termed Measure Gesture. Measure Gestures possess configurational similarities to the ‘precision grip’ gesture family (Kendon 2004, see also Winter, Perlman & Matlock 2013). They are Holding gestures\(^75\) that profile specific articulators, that is specific parts of the hand: the index finger and thumb (distal phalanges); more precisely the finger pads, which we use to grasp small objects and explore objects by touching. Of the finger pads, the small tactile surfaces are profiled. The object is enclosed by these two opposed surfaces. They partially, only very incompletely, enclose the object, spanning (and thus disclosing) its height. The resulting form is a three-dimensional object that is largely undefined. It is specified merely in its (one-dimensional) vertical extent. Other Measure Gestures may specify horizontal or sagittal extents.

Many, if not all, shapes can be described by (for example vertical) extent, which is why in natural elicitation, Measure Gestures are usually accompanied by other gestures and language about space, both further defining the object’s shape. Figure 10.3 exemplifies shapes that this

---

\(^{75}\) For gesture types, only the name of the type is capitalised, in a specific group of gestures, also the word ‘Gesture’ is capitalised (‘Measure Gesture’).
gesture, on its own, could refer to: not only a cube (a), but also a small stick figure (b) or many other forms could be characterised in their height by this gesture. As long as the vertical extent matches the measured height, any shape is acceptable. Even empty space between two entities such as the ceiling height of rooms in a doll house (c) could be meant. (d-f) thus clearly also fall into the category of potential referents of Measure Gestures. Their vertical diameters match the depicted height.

![Figure 10.3(a-f): Possible referent shapes of a Measure Gesture.](image)

According to Gesture Form Analysis, spatial information specified in Measure Gestures is very limited. Imitating the act of holding, this gesture can grasp any (three-dimensional) object. Still, the only specified measure is one-dimensional, as indicated by dotted red lines (Figure 10.2 and Figure 10.3). Thus, Measure Gestures can be, alternatively, conceptualised in one-dimensional space as a one-dimensional measure between two enclosing points, exemplifying the affordance of enclosure by two points in one-dimensional space (on the right of Figure 7.4).

**10.3.2. Shape Gesture**

Shape Gestures also hold an object, but they profile the whole index finger and thumb (and part of the hand between the two). In contrast to Measure Gestures, the profiled shape is one single surface curved from the tip of the index finger over the hand to the tip of the thumb. This regularly curved surface is then continued so as to fill up the small distance between index and thumb to a full hollow cylinder. This shape can enclose a three-dimensional object not only at the top and bottom, as in a Measure Gesture, but also from left and right. It more importantly conveys continuous curvature (Figure 10.4). At least in a vertical cutout plane, the object has to be circular (e.g., the sphere depicted in Figure 10.4). Thus, a Shape Gesture partially conveys information on shape (curvature) as well as size (height and width).
Figure 10.4: Shape Gesture, one group of gestures within the gesture type Holding.

The gesture form analysed in Figure 10.4 could alternatively be conceptualised in two-dimensional space (a curved line enclosing 2D space as in Figure 7.5). The constituent forms in two-dimensional space play out to produce a similar resulting form. The difference between a two- and a three-dimensional-space interpretation is not relevant for the hypothesis of this study, which only involves the distinction between curvature and no curvature. The same is true for the ambiguity of also the other fingers being profiled, as a hand shape as the one shown in Figure 10.4 could be interpreted. This interpretation would result in a longer cylindrical shape, but not in the loss of curvature information. All these interpretations show enclosure of a curved shape, either a curved line in two-dimensional space or a curved surface in three-dimensional space (on the right of Figure 7.5 and Figure 7.6, respectively).

A Gesture Form Analysis of a Shape Gesture interprets it to convey additional spatial constraints that narrow down the range of specific object shapes this gesture can refer to. The differently rounded shapes in Figure 10.5(d-f) are acceptable, while, in contrast to a Measure Gesture, the shapes in Figure 10.5(a-c) are not.

Figure 10.5(a-f): Possible referent shapes of a Shape Gesture.
While both Measure and Shape Gestures allow an observer to infer a three-dimensional shape, the degree and kind of specification of this shape show substantial differences. If it was only for the specified shape features, the spatial information of a Measure Gesture can be captured in one-dimensional space, while the Shape Gesture’s curvature information requires at least a two-dimensional space.

10.4. Attaining Measure and Shape Gesture interpretations

In trying to operationalize an observer’s perception, we face the general difficulty of semantics: we cannot know exactly what people think. Still, Gesture Form Analysis claims specific forms to be a necessary part of gesture conceptualisation. Consequently, if participants report that one group of gestures conveys spatial information that necessarily includes a specific gesture form and that another group of gestures conveys spatial information that necessarily includes a different gesture form, the participants’ perception of these two gestures is attainable (with regard to the specifics differentiating the two structures of gesture form). Attaining specific aspects of a participant’s gesture form interpretation allows for a semantic comparison, since the meaning of both Measure and Shape Gestures comprises specific kinds of spatial information. The occurrence of these different form interpretations could then be tested for statistically significant differences in articulator form.

It follows from the Gesture Form Analyses of the preceding section that both Measure and Shape Gestures convey enough spatial information to infer the size of the object. Note that neither of the two does so by virtue of the size of the hand or the fingers themselves, but by the space between the fingers that can vary from one gesture of a group to another gesture of the same group. Since size is conveyed in both groups of gestures, it is not suitable to differentiate between the two. More interestingly, Measure Gestures lack curvature information which is an integral part of Shape Gestures. This does not mean that Measure Gestures convey that the object is not curved. Information on whether the reference object is round or not is simply not part of what this gesture communicates. Measure Gestures are open to round, straight, zigzagged or any other kind of object shape (see Figure 10.3).

10.4.1. Interpretation by participants

To attain whether a participant sees (any kind of) curvature being conveyed by a specific, non-moving hand shape presented on a screen, I suggest the following question: “Can one obtain curvature information from what the hands are doing here?” („Kann man dem was die Hände
hier tun eine Rundung entnehmen?”). If the selection of the data successfully excluded everything that is not Measure or Shape Gesture, this answer distinguishes Measure from Shape Gestures as Figure 10.6 summarises.

<table>
<thead>
<tr>
<th>Observation \ Gesture</th>
<th>Measure Gesture</th>
<th>Shape Gesture</th>
</tr>
</thead>
<tbody>
<tr>
<td>Curvature</td>
<td>No</td>
<td>Yes</td>
</tr>
</tbody>
</table>

Figure 10.6: Curvature information deciding over participants perceiving Measure or Shape Gesture.

The observer task includes guiding the participant visually to clarify their decision on curvature. Figure 9.8 shows different kinds of curvature, of which only circular (a) or cylindrical curvature (b) are specified in a Shape Gesture. Circular curvature would be the result of neglecting the comparably little depth of the articulator unit thumb and index finger or of a conceptualisation in two-dimensional space, a cylindrical curvature would be the result of an analysis in all three dimensions as analysed in Figure 10.4. The Shape Gesture’s ambiguity between circular and cylindrical curvature is thus not relevant for the clear-cut separation from Measure Gestures since Gesture Form Analysis shows that Measure Gestures do not convey any type of curvature whatsoever and that Shape Gestures obligatorily convey either circular or cylindrical curvature, or both.

Some participants mentioned that any cylindrical shape must also include circular curvature, which is true (spherical shapes also necessarily include circular curvature). The tasks are designed to also be open to participants who are not aware of this geometrical hierarchy. These participants could for example select only cylindrical curvature, a response that is conflated in post-processing to the more general information that the participant identified (any) curvature to be conveyed in a gesture.

10.4.2. Interpretation by raters

Section 9.8.2 addressed the benefit of asking the participants to interpret their own gestures shortly after performing them. Gesture Form Analysis was used to argue a causal dependency between the identification of curvature and the classification as Measure or Shape Gestures. Still, independent empirical support of this theory-driven conclusion is desirable. To provide this, the author and two other expert raters independently rated all selected gestures as to whether they convey curvature, or not.
Raters not only supply further observations for cross-individual agreement, they are also a different kind of observer. In comparison to the participants of Study 2, the raters did not just perform the gesture that they interpret and they did not complete the tasks. Still, in this case, they did know of the objects, the tasks, and the design of Study 2. Thus, they are not purely naïve, independent observers, either. As a follow-up, participants of the extreme observer-end of the spectrum, truly naïve observers, could be asked to interpret curvature. These participants would have to be unaware of any context, instead just watching hand-shape stills. Positive results of participants of Study 2 and raters made this dispensable for the purposes of this dissertation.

The ratings are used in two ways: Inter-rater agreement is needed to assess whether different individuals reliably agree on one gesture being either a Measure or Shape Gesture. If there is high agreement between raters, the three separate ratings can be merged to one rating that reflects the majority of the raters and serves as (an approximation of) the correct interpretation of Measure and Shape Gestures for subsequent quantitative testing.

10.5. Articulator form differences between Measure and Shape Gestures

While the preceding section documented the assessment of systematic differences in gesture form, it did not address which visible physical cues favour a certain gesture to be interpreted as a Measure or a Shape Gesture. The next sections (until the hypothesis) document the search for details in articulator form that Sowa’s analysis did not consider, but which I claim to be systematically different in Measure and Shape Gestures. Once operationalised, these differences can be framed into a hypothesis for quantitative testing.

10.5.1. Predicted differences

Starting point for the development of the hypotheses was selecting all Measure and Shape Gestures of the single-participant video stimuli of Study 1. 14 Measure Gestures (Figure 10.7) and three Shape Gestures (Figure 10.8) were recognised. In accordance with earlier casual observations, certain characteristics of articulator form seemed to occur in Measure but not in Shape Gestures. Preliminary analyses of the other participants of Study 1 supported these differences. Two characteristics tended to indicate Measure Gestures:

1. Index finger being more extended (less curved).
2. Pinkie being more curled in (less raised).
Figure 10.7 and Figure 10.8 give an impression of Measure Gestures involving a more curled-in pinkie and less curved index finger, while Shape Gestures feature a less curled-in pinkie and a more curved index. Reasons explaining the choice of these specific aspects of articulator form are given below.

![Figure 10.7: Screenshots of the expressive phase of 14 Measure Gestures (Study 1).](image)

![Figure 10.8: Screenshots of three Shape Gestures (Study 1). The data of this participant (taken to elaborate Study 2) did not include more Shape Gestures (for more examples of Measure and Shape Gestures, see Appendix 7).](image)

Gesture Form Analysis claims that Measure Gestures are conceptualised as two small surfaces that oppose each other. The index finger should thus be straight so that the surface at the finger pad can be parallel to that of the thumb. Especially in the first ten pictures of Figure 10.7, which convey a small and medium-sized measure, this pattern is visible to the naked eye. By contrast, for example, the Shape Gesture in the middle picture of Figure 10.8 shows the index well curved despite index and thumb being close together. This exhibits a round inner
contour of the index-hand-thumb articulator unit profiled in Shape Gestures. The curvature of
the index is a cue for Shape Gestures (see Section 7.2.10 on cues of different alternative
interpretations).

The second cue, the extent to which the pinkie is curled in, might be surprising at first
sight. Since the pinkie is not part of the profiled articulator unit, why does it influence gesture
form? Does Gesture Form Analysis not suggest the pinkie to be not profiled for the meaning of
this gesture? It is true that the pinkie is neither relevant in the Measure nor the Shape
Gestures. On the contrary: this finger potentially disturbs the presentation of spatial
information—it is an obstacle, it occludes relevant spatial information. In Measure Gestures,
the extent between index finger and thumb is well visible even when the pinkie is curled in,
since the finger is not in the way of an observer computing the distance between the two
surfaces (Figure 10.7). In Shape Gestures, the pinkie would partially occlude the ‘C’-shaped
surface (Figure 10.8). Thus, I predict that a curled-in pinkie can be found significantly more in
Measure than in Shape Gestures. That is, if curl can be measured on a metric scale, the pinkie
in Measure Gestures should be more curled in than in Shape Gestures.

As Figure 10.7 and Figure 10.8 show, not only the pinkie, but also the middle and ring finger
occlude the ‘C’-shape of a Shape Gesture. This study focuses on the pinkie, since the middle
and ring finger often move with the pinkie (Section 9.7). A separate investigation of middle
and ring finger could add to the explanatory power of the tests conducted here.

10.5.2. Operationalising Curve

What do Curve and Curl mean? Answering this question requires defining variables that
capture these aspects of articulator form. Certain thresholds of these variables are to
differentiate Measure and Shape Gestures.

The Curve of a finger is understood here as the bending of the two intermediate joints
between the proximal, medial and distal phalanges. I regard only one variable as sufficient to
capture curvature for this investigation, since both joints seem to usually bend in a coupled
way during gesturing observed in Study 1 and 2.

The variable Curve is measured by the distance in centimetres between the two finger
markers: one at the tip, the other at the root of the finger (Figure 10.9). The distance is
inversely proportional to the curvature of the finger: The smaller the distance, the more curved
the finger.
Even if, in exceptional cases, the two joints are bent independently, this single variable will capture the bending of one single joint, since independent bending of one joint will also diminish the distance spanned by the finger. Since the research question does not focus on different kinds of curvature (e.g., circular vs. ellipsoid) but on whether participants interpret any curvature to be part of the gesturally conveyed information, or not, the variable of Curve can be captured by the distance between two points (Figure 10.9).

![Figure 10.9: The variable Curve inversely proportional to the distance spanned by the finger.](image)

Overstretching a finger could be seen as problematic for this variable, since an overstretched finger (that is curved backwards) is spuriously equated with a slightly curved finger. Figure 10.7, Figure 10.8 and non-systematic personal observation suggest, however, that overstretching does not occur in the index finger during the investigated gestures. This is not surprising, since Measure and Shape Gestures function, as analysed, curved inwards (index and thumb curving towards each other in a Shape Gesture) or parallel (Measure Gesture), but not being curved outwards.

10.5.3. **Operationalising Curl**

The pinkie can occlude the inner contour of index finger and thumb for an observer when it is not raised high enough. Figure 10.10 shows the pinkie, ring or middle finger partially occluding this contour. This contrasts to Figure 10.11, in which the pinkie is bent out of the way, leaving the regularly curved contour between index and thumb visible.
Both an only slightly curved pinkie (Figure 10.10a) and a more curved pinkie (Figure 10.10b,c) can occlude the ‘C’-shaped contour. Thus, Curl adds and conflates flapping a finger down at its root (joint between metacarpals and proximal phalanges) with the already defined Curve (see also Grandhi et al. 2012). Figure 10.12 and Figure 10.13 illustrate this conflation. The variable Curl differs from Curve in two essential aspects: it includes the joint between metacarpals and proximal phalanges and it is measured in degrees.

Figure 10.12 shows how Curve is calculated. On the basis of Hand Shape Coordinates (Section 9.9.1), the angle of the marker at the tip of the finger is calculated in relation to the up-down axis (horizontal line in Figure 10.11; y-axis in Figure 9.9). A finger stretched out in exact continuation of the hand has the angle of zero degrees, a partially flapped-down finger, as in Figure 10.12, a degree of roughly 35 degrees and an orthogonally inclined finger 90 degrees.
The angle measure as shown in Figure 10.12 and Figure 10.13 entails using Hand Shape Coordinates. Since the data contains the finger position divided into coordinates on three axes, one axis can be disregarded to obtain a profile view from a specific perspective. Calculating Curl, the in-out axis (x axis in Figure 9.9) is ignored. The result is a profile view as in Figure 10.12 and Figure 10.13, allowing to capture the angle as being projected on the plane defined by the up-down and back-front axes (plane running through the y and z axis). The wider the angle, the more curled-in the finger. Overstretching the finger is also represented by this metric variable, in form of negative values. In contrast to an overstretched index finger, an overstretched pinkie does occur (e.g., the pinkie Curl in the third-last picture in Figure 10.7 is around zero, potentially negative). The kind of angle measurement captured by the variable Curl can be imagined to be captured by a small camera mounted on the hand and looking onto the hand in the profile perspective shown in Figure 10.12 and Figure 10.13, no matter how the hand is rotated or where the hand is moved. Study 2 tests whether the specific aspects of articulator form, index Curve and pinkie Curl, significantly correlate with gesture form predictions for Measure and Shape Gestures.

10.6. Hypothesis

Gesture Form Analysis shows the entire curvature of the index-thumb articulator unit to be relevant in Shape Gestures but not in Measure Gestures. The following hypothesis is put forth:

The index finger in gesture holds interpreted as Shape Gestures is predicted to be more Curved than in Measure Gestures because, in Shape Gestures, it exhibits the curvature of the referent object. Also, the pinkie is predicted to be less Curled in during Shape Gestures than in Measure Gestures, to not occlude the 'C'-shaped curvature conveyed in Shape Gestures. Evidence for this hypothesis would support the existence of gesture form as non-equivalent to articulator form. It would also support the specific Gesture Form Analysis of the investigated gestures: two opposing surfaces in Measure Gestures and one curved surface in Shape Gestures.
10.7. **The overall workflow of Study 2 including manual data selection**

Before presenting the results, this section will describe the workflow including how relevant data is selected out of the raw video and motion-capture data. Measure and Shape Gestures were selected from the whole data recorded for Study 2. This section lays out the methodology of Study 2, combining automatic computations with manual data selection, building up on the introduced methods that include scripts and a software plugin as part of the data post-processing as described in Chapter 9.

The first step is the generation of all desired variables from the raw Motion Capture data by the Java script MoJa. The variables include the Hand Shape Coordinates on the basis of the Minimal Marker Set and various distance and angle measures, but also for example the speed of the marker at the tip of the index finger, which defines the speed of the hand.

A hand speed of less than ten centimetres per second, regardless of the direction of the movement, is a necessary condition for a gesture hold. The Elan Plugin for Automatic Annotation (EPAA) determines occurrence and duration of all holds in the entire data set by creating annotations in Elan. Certainly, not all low-movement phases in the participants’ hand movements are holds. But the automatically created annotations already contain the low-movement phases that are holds, including their objectively defined duration.

Within Elan, the non-relevant holds are manually deleted. The manifold criteria for the deletion of certain low-movement phases, for example when the hands are in the rest position, will be laid out in the following Section 10.7.1-5.

Once all the relevant Measure and Shape Gesture holds are selected, the Data Extractor computes variable averages (from MoJa-generated data) during the duration of the holds. This results in a list of all observations of Measure and Shape Gestures in Study 2. Each observation is assigned averages of variable values including those representing index finger Curve and pinkie Curl. These observations are then used for the statistical analysis.

The following sections state exclusion criteria that will leave only those low-movement phases that are (1) gesture holds and (2) either a Measure or a Shape Gesture.

10.7.1. **Excluding rest positions**

The data selector (one of the expert raters) also modified the automatically generated annotations in Elan following specific instructions.
The focus of this investigation is on holds, which are included in low-movement phases but are more specific. The hand has to be held still during, before or after the expressive phase of a gesture. All other low-movement phases, including the hand in rest position, are excluded from the study data (deleted within Elan) because they are not part of a gesture.

This deletion process is not necessarily a manual task that requires human interpretation. Rather, follow-up studies are intended to automatise these exclusion criteria, for example by deleting any low-movement phases that occur at the height of a seated participant’s lap (a variable directly provided by raw motion-capture data). The following criteria are also candidates for automation. Since this investigation’s time limitations did not leave room for intensive implementation and testing, these processes were carried out manually to assure coherent data.

10.7.2. Only index-and-thumb profile

Not just any gesture is included, but only gestures that involve a specific articulator unit. Measure and Shape Gestures belong to the group of Holding gestures. While there are several gestures that Hold an object, this investigation includes only those in which the index finger and thumb of one hand alone convey spatial information. Examples of Holding as in Figure 8.30 and Figure 8.31 do not do so and are thus excluded in the data selection.

10.7.3. Excluding gestures conveying spherical curvature

To limit the investigation to Measure and Shape Gestures, which are similar in articulator form, gestures conveying spherical curvature (according to the participants’ decision) were excluded. Gestures conveying complex spherical or related kinds of curvature information as in Figure 8.33 and Figure 8.39 are performed in a different way and are of a different articulator form.

While asking the participant for different kinds of curvature (Figure 9.8) mainly served to reassure whether the participant identified any kind of curvature at all, it also helped to exclude gestures conveying spherical curvature.

10.7.4. Excluding ambiguous gesture form

While this investigation focuses on holds, gesturing conveys spatial information by motion as well (as discussed at length, for example, in Section 7.2.9 for the operation of Trace Leaving and in Section 8.5.2 for the gesture type of Holding and moving). Participants could identify a
curvature in the movement trajectory that precedes or follows the hold, which would be erroneously categorised as a Shape Gesture. This amalgamation is problematic for precisely defining Shape Gestures on the basis of participant interpretations of curvature information alone.

To address this problem, data selection included watching the whole gestures, not only the holds. Whenever other parts of the gestural movement could be interpreted to convey curvature (e.g., a curved motion that could be interpreted as leaving a trace) this annotation is deleted, excluding this gesture as an observation for the test of the specific hypothesis about Measure and Shape Gestures.

10.7.5. The temporal relation of low-movement phases and holds

The investigation focusses on gesture holds. Low-movement phases are much shorter than what can be expected of a regular gesture-phase coding. Even a minor negative peak that is below the threshold speed for only one millisecond results in EPAA creating a new annotation in Elan. This problem is countered by the following lower limits to the duration of one annotation and the duration of a gap between two annotations.

The objective criterion of low-movement phases also resulted in EPAA creating large numbers of annotations, sometimes more than ten annotations in what the participant interpreted to be one gesture. Since any annotation will be one observation in the statistical testing, this method would produce many more independent observations than there actually are. To reduce the number of observations, the data selection includes deleting all very small annotations of 200 milliseconds or less.

Further, the automatically determined low-movement phases did not tolerate minor fluctuations in the speed of a gesture hold. When gesturers hold up their hand in a Shape Gesture configuration and only slightly superimpose a beat, for example to emphasise what they are just saying, two low-movement phases were generated because the hand speed exceeded ten centimetres per second for the very brief time interval of the rudimentary beat. To attenuate this source for an inflationary amount of observations, a gap between two annotations with a pause of 100 milliseconds or less is merged to one annotation.
<table>
<thead>
<tr>
<th>Order</th>
<th>Criterion</th>
<th>Threshold</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Upper speed limit, measured at the tip of the index finger</td>
<td>10 cm/sec</td>
</tr>
<tr>
<td>2</td>
<td>Lower limit of annotation length (deleted first)</td>
<td>200 ms</td>
</tr>
<tr>
<td>3</td>
<td>Lower limit of annotation pause (deleted second)</td>
<td>100 ms</td>
</tr>
</tbody>
</table>

Figure 10.14: List of quantitative exclusion criteria for the Measure and Shape Gesture annotations. The order in this list is also the order of execution.

10.8. Results

The processed data of Study 2 contained 471 observations (191 Measure, 280 Shape Gestures). On the basis of these data, the following sections report on inter-rater reliability ratings, the result of the hypothesis tests on the difference between Measure and Shape Gestures, and the distribution index Curve and pinkie Curl in Measure and Shape Gestures.

10.8.1. Difference between Measure and Shape Gestures confirmed

Figure 10.15 shows that all individuals who rated Study 2 tend to agree in their interpretation of Measure and Shape Gestures according to Cohen’s and Fleiss’ kappa. The agreement throughout all combinations of interpreting individuals is “substantial” (Landis & Koch 1977).

<table>
<thead>
<tr>
<th>Rater A</th>
<th>Rater B</th>
<th>Kappa</th>
</tr>
</thead>
<tbody>
<tr>
<td>Author</td>
<td>Rater 2</td>
<td>Cohen’s κ=0.802</td>
</tr>
<tr>
<td>Author</td>
<td>Rater 3</td>
<td>Cohen’s κ=0.805</td>
</tr>
<tr>
<td>Rater 2</td>
<td>Rater 3</td>
<td>Cohen’s κ=0.803</td>
</tr>
<tr>
<td>Author, Rater 2 and 3</td>
<td></td>
<td>Fleiss’ κ=0.803</td>
</tr>
<tr>
<td>Participants</td>
<td>Author</td>
<td>Cohen’s κ=0.728</td>
</tr>
<tr>
<td>Participants</td>
<td>Rater 2</td>
<td>Cohen’s κ=0.638</td>
</tr>
<tr>
<td>Participants</td>
<td>Rater 3</td>
<td>Cohen’s κ=0.621</td>
</tr>
<tr>
<td>Author, raters and participants</td>
<td></td>
<td>Fleiss’ κ=0.729</td>
</tr>
</tbody>
</table>

Figure 10.15: Cohen’s and Fleiss’ κ respectively for all combinations among the individuals interpreting Measure and Shape Gestures.

76 These thresholds were determined by discussing the output of several different velocities and time limits with members of the nmLab.
More importantly, the hypothesis was corroborated. The index finger was significantly more Curved in Shape Gestures and the pinkie significantly more Curled in Measure Gestures. Figure 10.16 shows the means of the defined variables.

<table>
<thead>
<tr>
<th></th>
<th>Measure Gestures (N=191)</th>
<th>Shape Gestures (N=280)</th>
<th>Two-sided t-test</th>
</tr>
</thead>
<tbody>
<tr>
<td>Index finger Curve</td>
<td>9.2 cm (SD = 1.1 cm)</td>
<td>8.7 cm (SD = 0.9 cm)</td>
<td>t(459) = 8.5777; p&lt;0.0001 ***</td>
</tr>
<tr>
<td>Pinkie Curl</td>
<td>83.1° (SD = 36.0°)</td>
<td>36.4° (SD = 31.6°)</td>
<td>t(373) = 14.498; p&lt;0.0001 ***</td>
</tr>
</tbody>
</table>

Figure 10.16: Significant differences in means of Measure and Shape Gestures.

These results show the variables to differ significantly in the expected direction. It is not surprising that the difference in index finger curvature is not very big, since moderate curving results only in minor distance changes. A significant result of half a centimetre difference of means is, in that light, consoling. The big difference in pinkie Curl is examined more closely in the following.

Figure 10.17 shows boxplot distributions of Curve and Curl in gestures that were predominantly rated to not convey curvature (Measure Gestures) and those that did convey curvature (Shape Gestures). While both differences in distribution are significant, the difference in pinkie Curl (Figure 10.17b) is very big. No data point within the margin of the 50 per cent most central angle values overlaps with those of Shape Gestures. This is coherent with the large difference in means and shows that the extent to which the pinkie is curled in is a valid cue for the differentiation of Measure and Shape Gestures. The upper outliers of Curl values for Shape Gestures (b) predominantly stem from one trial of one person who performs one very long gesture hold.
The scatterplot in Figure 10.18a shows the distribution of Measure Gestures (red triangles) and Shape Gestures (black circles) with regard to index Curve (y axis) and pinkie Curl (x axis). Scatterplot (b) shows the data to be roughly dividable into four quadrants, to which (c) maps exemplary hand shapes.

As predicted, straight index and curled pinkie occur in gestures predominantly rated as Measure Gestures as illustrated by the density of circles in the upper right quadrant of (b). Curved index and raised pinkie are often Shape Gestures as shown by the triangle-dominance in the lower left quadrant. While very few observations fall into the lower right quadrant, the upper left quadrant is occupied by a mixture of both gestures.

There is a mix of Measure and Shape Gesture interpretations in gestures with extended index and raised pinkie. I attribute this to the fact that both gestures’ articulator form become more similar for larger objects for physiognomic reasons. To gesturally represent the more stretched-out curvature of a larger object, the index finger cannot be as curved as it is for small objects. Thus, it does not differ as much from displaying a large one-dimensional measure. Not only the index finger but also the pinkie is more difficult to differentiate for larger objects. Even in Measure Gestures, the larger the object, the easier a Curled pinkie can get in the way of the space between index and thumb. In Study 1, the participant also raised the pinkie considerably in Measure Gestures conveying large measures (last three pictures of Figure 10.7).
Figure 10.18: (a) Scatterplot matrices of index Curve (distCm.Ah.pi_i) and pinkie Curl (angleDeg.Ah.pp.p), (b) superimposed by lines indicating the quadrants and by (c) example gestures of three configurations.
10.8.2. Categorical difference between Measure and Shape Gestures?

Another distribution demands further investigation. When the proportion of Measure to Shape Gestures is plotted against the pinkie Curl, a continuous increase of Shape Gestures could be expected. Figure 10.19 shows this not to be the case. Rather, we see stagnating (and temporarily decreasing) values at the extremes of pinkie Curl and a steep rise in the middle. The resulting curve shows similarity to a sigmoid function—which is taken as evidence for categorical perception (Liberman et al. 1957; McMurray & Spivey 1999), differentiating for example consonants like “ba” versus “pa” by voice-onset-times (Pisoni & Tash 1974). The steep rise is taken as evidence for the perception of two distinct categories of sounds: phonemes. Similarly, Gesture Form Analysis claims Measure and Shape Gestures to convey different categories of information (a measure vs. a whole shape contour). This finding needs to be evaluated carefully with respect to the differences of the modalities of speech and gesture. Some first thought on this evaluation are put forth in Section 11.3.

Figure 10.19: Frequency of Shape Gesture interpretations in 10 stages of decreasing pinkie Curl (low x-values more curled, high x values less curled).
11. Discussion

This dissertation dealt with a model that captures necessary aspects of spatial gesture interpretation as part of basic semantic processing. Chapter 3 and 4 proposed surfaces to be of particular importance in gesture conceptualisation. As part of specific Gesture Form Analyses, two gestures were claimed to conceptually involve different profiled surfaces. The model received empirical support by the hand shapes of these two specific gestures, which indeed behaved as predicted regarding the extent and position of these surfaces.

11.1. Characteristics of Gesture Form Analysis

Study 2 was based on Gesture Form Analyses of two gestures whose physical configuration is very similar, but whose gesture form is markedly distinct: Measure and Shape Gestures. Gesture interpretations agreed reliably across four individuals (three raters and the respective participant), enabling an investigation of gesture form differences between the two gestures. Analysis of motion-capture data showed significant configurational differences in gestures that convey a single size measurement (Measure Gestures) versus those that convey the contour of a shape (Shape Gestures). For example, the pinkie was raised in order to not occlude the imaginary shape which Gesture Form Analysis identified as compulsory: the ‘C’-contour of a Shape Gesture.

For Measure and Shape Gestures, this is evidence for manual action adapting to imaginary forms proposed in Gesture Form Analysis. The evidence is confined to one distinction among two specific gestures. Since this distinction bases on the profiled surfaces proposed by Sowa (2006) and in Gesture Form Analysis, the evidence is interpreted to be in support of the perspective taken in these approaches.

The rest of the proposed typology was developed based on systematic analysis of a range of gestures and it may be adapted to other contexts of gesturing in further studies. Future adaptations might lead to new combinations of gesture form operations and thus changes to different branches of the gesture typology (Figure 8.59). The core goal of this dissertation is to introduce a new approach to gesture form, which includes imaginary forms as well as physically present forms. It was demonstrated that gesture form is necessary for understanding gestures. Not only the resulting form is necessary, but also each of the constituent forms that a gesture is made up of—at least as intermediate steps in order to compute the resulting form.

In various cases, the concept of these constituent forms is already implicit in the literature. By means of a simple example, I believe it is undisputed that an observer of someone pointing at a
distant object has to continue the direction indicated by the pointing finger in order to encounter the object. To these and many other of the spatial relations observed in gesture analyses this dissertation did not contribute any entirely new—let alone competing—claims (including, but not limited to, internal and external metonymy (Mittelberg), types of sketching (Mandel), constraints of descriptive gestures (Sowa), constraints of pointing gestures (Fricke; Kita; Wilkins), the hand as different tools (Müller; Streeck)). The proposal of this dissertation is, in a first step, to delimit operations as for example the operation of Continuation from other spatial operations necessary for certain gesture types and specify this operation’s geometrical constraints. In a second step, collect all the spatial operations and combination of operations that could be found in Study 1 and 2 and in the literature. Finally, all the different combinations of operations (gesture types) were listed with their constituent in a gesture typology.

I would like to reaffirm that Gesture Form Analysis does not challenge Stokoe’s (2005/1960) form parameters or any other method for describing physical movement. Rather, it is seen as an addition to the physical analysis and it roots in the affordances of different kinds of form: a point, a line, a surface and a volume (see Chapter 7). While the form parameters are productive in distinguishing a variety of functions a gesture can have (see Kendon 2004: Chapter 11, among many others), they are not always constant within one gesture type. Pointing At Object can be performed with different hand shapes, motions, and at different locations in gesture space. Thus, neither a single parameter nor a combination of parameters could select pointing gestures out of a large set of gestures—or the list of parameters would have to be very long, including all the ways in which different articulators can point. As can other functional classifications, Gesture Form Analysis allows selecting all pointing gestures of a data set, or a specific type of pointing gesture. In contrast to existing functional classifications, Gesture Form Analysis is modular and strictly form-based. It is strictly form-based in that conceptual primitives of form (form dimensions) define the structure of the analysis. Since a form of a certain dimension affords specific functions, functional distinctions can be broken down to form distinctions. It is modular in that a small set of spatial operations (eight gesture form operations) account for all the constituent forms of all gesture types encountered in the studies and in gesture literature. Every operation is defined by how it creates form. More specifically, on the basis of an input form, every operation constraints the dimensionality of the output form, as expressed in a simple formula (Chapter 7). Gesture form, as defined in this dissertation, is claimed to be a fundamental structure in the cognitive process of gesture conceptualisation.
11.2. First thoughts on the integration into Talmy’s recent work

Speech, the spatial surrounding and various other aspects of context were mentioned in the specific analysis (Chapter 8), but did not obtain focus or systematic distinction. This was to highlight what communicative content gesture alone can convey, since I believe this perspective has received only comparatively little attention. Yet, different kinds of context are crucial in naturalistic gesturing, as Goodwin (e.g., 2003a), Hutchins (e.g., 2005) and many others show.

In this way, Gesture Form Analysis requires not only adaption to further data and testing of the proposed distinctions, it also needs further systematic integration in the conceptualisation of an utterance as a whole, including the different aspects of context. A systematic investigation of how an interaction of gesture form and other factors could look like is part of recent work by Talmy (forthcoming), focusing on pointing gestures or “targeting” gestures. Talmy identifies the spatial information conveyed in gesture as belonging to one of ten “cue categories” (Talmy 2012) that helps eliminating other possible targets in order to direct the addressee’s attention to the desired target. For example with the help of “lexical cues” (saying “these” to refer to a plural, proximal target), “chronal cues” (saying “this bathroom is free now” while pointing to a bathroom which is declared to be free in the moment of the utterance), epistemic cues (saying “this is my father” pointing at a pair of persons of which only one is of an appropriate age to actually be the father), the possible targets of a pointing gesture are systematically disambiguated—in addition to the focus afforded by the gesture (or ‘gesture form’, as I would like to specify).

As a first small step towards integration, the term ‘cue’ was used in this dissertation in an adapted definition. Talmy uses ‘cue’ as an indication favouring one of many candidate targets. This definition was extended to those gestures that do not target but nevertheless need clues in order to interpret their form (Chapter 8). Thus ‘cue’ was used here as an indication favouring one of the theoretically possible gesture form interpretations, in any gesture type, as for example in favouring either a Measure or a Shape Gesture interpretation of an ‘index finger and thumb extended’ gesture.

11.3. Two coarse speech analogies and the problem of categorical perception

Does gesture form structure gestural action similar to how grammar structures language? Like Müller’s modes of representation (1998; 2010; 2014) as well as Müller and colleagues’ work “towards a grammar of gestures” in “a form-based view” (Müller et al. 2013), this dissertation
and the notion of gesture form as a systematic aspect of gesture conceptualisation could be seen as akin to grammar in linguistics. Parallels, for example topological distinction of form as described by Talmy (2000:25-31; see Chapter 5), further justify the search for a grammar of language and gesture (see also Fricke 2013).

Despite discussing the analogy of grammar here—as in the possibility of a structure of gesture that is closely related to the way grammar structures language—the aim of this dissertation was not to work towards a grammar of gesture. It would be surprising if structures within the motor-visual modality of gesture could generally be described by the structures in the aural-oral speech modality. Rather, the model of gesture form presented in this dissertation was fuelled by the spatial structures found in gesture studies, Sign Linguistics, and perception of space by vision and touch. They are rooted in the structure of human (manual) action on objects, in a phenomenological understanding (Arnheim; Katz; see Chapter 4).

The potential side effect on the categorical perception of gestures in Study 2 (Section 11.3) has to be seen in this modality-specific perspective. Saying that Measure and Shape Gestures are distinct categories with one prototype hand shape each is not an acceptable prediction, because it ignores the structural similarity promoted in Gesture Form Analysis. The hand shape and movement are not similar within one gesture type, but gestures of one type are perceived as imitating the same practice—despite the variability whose constraints were investigated here and in Sowa (2006). Thus, the possibility of a categorical distinction akin to categorical perception in speech has to be rejected. Participants do not simply recognise, for example, a hand shape resembling what they learned a Shape Gesture hand shape should look like. Tokens of gestures introduced as Shape Gestures can enclose shapes of different sizes and different curvatures. Thus different hand shapes are naturally expected even within the small group of Shape Gestures (very specific sub-category of Holding gestures). Observers rather identify the action that this gesture imitates: holding something in the whole bend of index finger and thumb. They imagine the gesture Holding an unknown |object| in a specific way. While knowing very little about this |object|, observers do know from their experience, that the |object|’s shape should fit into the inner surface of index and thumb, and this brings us to an acceptable explanation for the rather sharp rise in Shape Gesture interpretations in Figure 10.19. There is a finger angle, where the pinkie moves in front of the curved contour of a Shape Gesture (or at least disturbs the clear ‘C’-shaped contour in a profile view). Precisely at this angle, the presentation of the imaginary ‘C’-curved surface in a Shape Gesture begins to be hampered. Since every Shape Gesture has a different curvature and every observer looks at the gesture from a different perspective, this angle varies. The point where the occlusion begins
seems to vary within the range of Figure 10.19’s steep rise. Gesture perception is thus not simply categorical with regard to one parameter but sensitive to gesture form as an interaction of articulators and observer perspective within the spatial context of the gesture. Controlling for the exact angle of an observer’s perspective and exact information on finger configuration could demarcate the border between Measure and Shape Gestures, allowing for a refined claim on perception in discrete categories on the basis of Gesture Form Analysis.

In the case of Pinkie Curl, also the non-profiled fingers can completely change the conceptualisation of a hand shape. They interrupt the regularly-cut contour that is central to the perception of a Shape Gesture. When the pinkie occludes the ‘C’ shape, the Shape Gesture interpretation is rejected, or at least becomes less likely. The Measure Gesture interpretation is untouched by the curled-in pinkie, thus becoming a more likely interpretation. To merely call this effect categorical is not specific enough. What it shares with the concept of categorical perception in speech is merely the categorical switch between two interpretations. To pinpoint the locations of all multifactorial borders of one group of gesture (let alone one entire gesture type or all gesture types) is a task for which Gesture Form Analysis, and the methodology of Study 2 provides only a first small step.

As touched upon in the Introduction, another coarse analogy connecting the domains of gesture and speech may be discussed. (1) Articulator form describes the physical phenomenon of visible articulators moving in space resembling phonetics that treat sounds as the audible phenomenon of speech. (2) The constituents of gesture form differentiate gesture types resembling phonemes that are differentiating meaning, but are not meaningful on their own. (3) Constituent forms combine to meaningful gesture form resembling phonemes combining to morphemes, the meaningful elements in speech (Figure 11.1). Again, this analogy is not to map linguistic structure onto gesture. It more generally sketches analytic layers of two forms of natural communication in which the physical phenomenon is processed by meaning-differentiating elements which combine to meaningful units (as shown on the right side of the typology, Figure 8.59).\(^77\) Attempts to differentiate movement units and map them onto meaningful or meaning-differentiating units go back to Birdwhistell, who termed these units “prekinesics”, “kinics” and “kinemorphics” (Birdwhistell 1970:198). Though not adopting his terminology, this dissertation’s focus on the gesture modality of body movement reflects Birdwhistell’s kinesics perspective.

\(^{77}\) I thank Cornelia Müller for her input on this analogy (personal communication).
The typology at the end of Chapter 8 is not seen as a closed system, it is rather just a framework for the classification of gestures. Neither the typology nor the gesture form operations go into detail about specific computations of object shapes in specific gestures, as did Sowa (2006). The goal of this thesis was simply to propose a methodology for gesture semantics by providing (1) the necessary elements underlying spatial processing, and (2) a basic structure of dimensions and operations for decomposing gesture types. The analyses at hand go back to Goffman’s perspective on interaction that has a clear focus upon the principles of interaction itself, not only “upon the outcomes of interaction. Goffman’s [focus] was upon how it was done; upon how, indeed, it was possible at all” (Kendon 1988b:19). As exemplified for Directing (Section 8.3.2) and Holding (Section 8.4.2), each gesture form operation is necessary for a specific gesture type to fulfil its function (indicating a direction or showing the size of the object held between index and thumb). This approach then is to stress the importance of spatial relations in a modality for which space is a Condition of Possibility (Kant 1787). Further, it is to highlight the affordances of abstracted zero-, one- and two-dimensional forms that are key in structuring spatial content—particularly in gesture. The dissertation also reiterates the demand to investigate “the embodied process of bringing objects and events into quasi-presentation” in interaction (Nemirowsky, Kelton & Rhodehamel 2012:130; see also Mittelberg 2013; Parrill & Sweetser 2004), for example to “measure’ (thumb and index finger stretched apart)” an object’s size (Mittelberg 2013:359).
12. Conclusion with suggestions for gesture form’s many loose ends

Gesture form, as introduced in this dissertation, is a model based on assumptions whose validity is subject to the reader’s evaluation and further research. The roots for these assumptions lie in a Cognitive Linguistic perspective on gesture, applying semiotic relations to the spatial modality of gesture. In this perspective, reformulating semiotic relations as specific geometric operations for gesture analysis is possible and, I believe, productive. Work on metaphor (Müller) and metonymy (Mittelberg) was cited earlier as the foundation for some of the operations developed here. More generally, the description of gesture form advocated in this dissertation is immediate support that, “metaphor [and, for gesture form, even more so metonymy] is pervasive in everyday life, not just in language but in thought and action” (Lakoff & Johnson 1980, italics added). The study of gesture—a kind of action (Wilcox 2004), a “window into thought” (McNeill 2000, italics added) and natural part of language use (“part of what the man was trying to say”, Kendon 2004:11)—is particularly apt for such kind of support. The novelty of this dissertation is the focus on gesture form, suggesting it as the basis of gesture semantics. Since gesture semantics is a very young field of interest much less investigated than linguistic semantics, I would like to emphasise that the proposed gesture form operations are minimalistic. Yet, the framework presented in Gesture Form Analyses of different gesture types is open to empirical investigation, including quantitative hypothesis testing. For this reason, Part II of this dissertation specified the necessary spatial operations in the form of geometric equations (Chapter 7) and assigned these spatial operations to the investigated gestures (Chapter 8). In this way, further research can take any Gesture Form Analysis as a claim that the specified geometrical operations are part of the conceptualisation of the respective gesture type. Thus, aspects of the theory can readily be empirically falsified. This is not always a given in work on gestures and Cognitive Linguistics.

This dissertation does not provide specific conclusions apart from the stated theory, methods and one piece of evidence. Instead, its theoretical model offers many loose ends for more empirical investigation. The understanding of gesture form being non-equivalent to articulator form marks a research gap. It is difficult to foresee, which, if any, consequences indicating this gap may have, yet I would like to close by proposing further studies and specific elaborations of the model developed here, which show this theory’s possible use.

There are several ways to expand the explanatory potential of the concept of gesture form, including immediate follow-ups on Study 2. Gesture Form Analysis also provides a basis for
developing a new perspective on gesture within the interdisciplinary field of gesture studies, including specifically Computer Science and Sign Linguistics.

12.1. Potential follow-up studies

As an immediate follow-up on Study 2, a naïve-observer study is planned (Study 3). A large number of participants who have no knowledge about the context of Study 2 are to rate whether gestures presented in the form of screen shots convey curvature or not. Stimuli for Study 3 would span the whole variety of index Curve and pinkie Curl, presented in a randomised fashion. If participants began to see a Measure Gesture in those Curve and Curl values at which the pinkie crosses the index-finger-thumb contour (claimed to be part of the gesture form of a Shape Gesture), this would strengthen the results presented in this dissertation, providing support for the context of the lab experiment not being necessary for the interpretation of forms as predicted by Gesture Form Analysis. A set of studies including different stimuli taken from natural gesturing in various contexts and precisely controlled computer-generated hand configurations could test claims of universality that were deliberately backgrounded in this dissertation. In experiments with larger sample sizes, Hassemer & Winter (forthcoming) report on a significant difference between Measure and Shape Gestures. Both in the gesture performance of pedestrians in a street experiment as well as in the perception of a continuum of computer-rendered hand shapes in an online experiment, more curled in middle, ring and little fingers occurred more often in Measure Gestures.

Furthermore, the whole spectrum of psychological experimentation including reaction time experiments or neuro-imaging techniques could provide further methodological means for more controlled follow-up experiments of gesture form. Here are two hypotheses for different neuro-imaging devices:

One hypothesis that could be derived from the grammar analogy of gesture form operations would be to test for activations associated with grammatical errors. In brain activity measured by electroencephalography, a peak 600 milliseconds after the production of the spoken word is associated with the processing of wrong grammar (the so-called ‘P600’; see, i.a., Hagoort 2007). Thus, violations of the constraints proposed by Gesture Form Analysis could be used for such experiments. If, analogous to grammatical mistakes in spoken or signed languages, a violation within a movement pattern produces a positive deflection associated with grammatical errors or syntactic anomalies, this would be evidence for a central, grammar-like role of gesture form. While the P600 is characteristic of signed and spoken language
perception (Capek et al. 2009; Liu et al. 2011), it also occurs in “harmonically inappropriate chords” (Maess et al. 2001), motivating experiments that investigate whether the structure of gesture belongs to the group of signals that are rule-governed in the way that full-blown languages and musical harmony are.

With regard to visual perception, studies involving patients with associative agnosia could further test the nature of different constituent forms in a single gesture. According to Farah (1991), there are two distinguishable abilities that can separately be impaired in agnostic patients. The first is difficulties in perceiving the shape of a single complex object, for example, a face. The second is difficulties capturing the shape of multiple parts that make up a composed object, for example a word being composed of several letters.

If we assume that our ability to know whether a single part has come from a particular object is roughly indicative of whether that part is explicitly represented in our structural description of the object, then we can hazard some guesses about which objects are recognised via decomposition into simpler parts and which are recognised with little or no decomposition. Face recognition seems to be a likely candidate for the latter type of process.

Farah 1991:376

Gesture Form Analysis claims Representing in three dimensions to consist of only one complex constituent form and Pointing At Object or Surface Directing to consist of multiple forms or parts. This would predict participants with face recognition impairments to also have difficulties in interpreting Representing (in 3D), and participants with word recognition impairments to also have difficulties in interpreting complex pointing gestures.

Any of the mentioned further directions of study could be tested in different populations. For example, cross-cultural and cross-linguistic studies would test to which extent a model of gesture form is culture or language-specific—or universal.

12.2. Gesture Form Analysis coding scheme

In an attempt to be thorough, Gesture Form Analysis as presented in this dissertation is maximally inclusive with regard to alternative interpretations. Since our knowledge of gesture semantics (independent of speech) is very limited, the qualitative analysis in this dissertation was designed to make room for all conceptualisations that cannot be excluded according to pertinent cues. The detailed analysis thus always includes considering possible alternative
interpretations, each demanding new analyses and possibly producing different resulting forms.

Gesture Form Analysis could be developed into a quantitative coding scheme similar to Lausberg’s Neurology-oriented NEUROGES system (Lausberg & Sloetjes 2009). NEUROGES is characterised by clear-cut hierarchical categories of movement and function that allow for a stringent workflow and interrater agreement tests. A hierarchical coding scheme of Lausberg’s type cannot make room for alternatives. Gesture Form Analysis also would have to be reduced to the dominant interpretation per gesture, streamlining the coding procedure.

The data generated by applying such a coding scheme would provide many tokens of certain gesture types. Combined with motion-capture data, any correlations within and across the types could be investigated, giving a picture of the variability within types and the differences and similarities between the types, potentially providing empirical scrutiny of the gesture types analysed in Chapter 8. Computational methods, as the one touched upon in the next section, could be used to interpret such data.

12.3. Contribution to computer-aided studies and human-computer interaction

Study 2 was restricted to two simple cues in two very specific gestures. Since it is difficult to find specific correlations in movement aspects of an articulator unit as complex as the hands (by looking at the video as well as by looking at the positions of motion-capture markers), machine learning might be a way to find candidates for significant cues. In a nutshell, machine learning is the search for continuous regions (clusters) in n-dimensional space, in which every dimension is one variable. The variables used in this study, including all Hand Shape Coordinates, Curve, Curl and raw motion-capture coordinate data are candidates for the dimensions of this ‘variable-space’. Such a computer-aided investigation opens up possibilities for Gesture Form Analysis in communication involving computers (human-computer interaction).

For machine learning to produce meaningful results, the proposed gesture typology (Section 8.8) has to be developed further. The typology should capture all recurrent gesture types in data collected in different contexts. On the basis of such a typology, computer-aided methods, such as machine learning, could collect cues for all gesture types. A collection of significant cues per gesture type could even enable automatic recognition. Such a development might be interesting for scholars in the field of Human-Computer Interaction. It is surprising
that, today, hardly any touchless gesture control is marketable (with the exception of game consoles). This is in contrast to the great tracking abilities of affordable hand and finger trackers (such as the device provided by the company LEAP). Gesture Form Analyses provides a typology of comparable gesture types that could provide the parameters needed to render gesture form identification better tractable. Likewise, Gesture Form Analysis may contribute to gesture synthesis in virtual animated agents (i.a., Bergmann, Kopp & Eyssel 2010; Kopp, Sowa & Wachsmuth 2004).

Moreover, the immediate touchless control of gesture interpretation systems could profit from knowledge about the spatial relation of the hand and the object. Since “people are generally able to interpret the transitive action as well as characteristics of the object manipulated despite individual variations in how people naturally gesture” (Grandhi et al. 2012:1), gestures involving imaginary objects seem to be equally suitable to control machines as non-transitive gestures. In the search for natural gestures that are machine-identifiable, Gesture Form Analysis could offer further differentiation on the kinds of spatial information gestures convey, especially those enclosing the imaginary object.

Since the emergence of digital motion-capturing in the late 1970s, the performance of motion-capture systems has continuously advanced. There is no indication that this trend is going to attenuate in the future—on the contrary, it is likely that precise tracking of body movements will become better and more affordable. The easier it becomes to capture articulator form, the more apparent is the need for a model of gesture form. In other words, the more the technology advances, the more will the lack of knowledge between what a gesture looks like and what a gesture means be relevant for the field of human-computer interaction studies and the industry.
Declaration of Authorship

I hereby certify that this thesis has been composed by me and is based on my own work, unless stated otherwise. No other person’s work has been used without due acknowledgement in this thesis. All references and verbatim extracts have been quoted, and all sources of information, including graphs and data sets, have been specifically acknowledged.

Berlin, 5th of May, 2016
References


Sutherland, N. S. (1960). The methods and findings of experiments on the visual discrimination of shape by animals. Cambridge: Heffer.


Appendices (Study 2)

1. Ethics committee approval

Studies 1 and 2 were approved by the Ethics committee at the Medical Faculty of the RWTH Aachen University, Pauwelsstraße 30, D-52074 Aachen as part of the ethics approval EK 115/10 “Quantitative analyses of coverbal gestures.” Investigators: i.a., Prof. Irene Mittelberg, Ph.D. (main investigator), Julius Hassemer.

2. Open access policy

Since this dissertation uses motion-capture data, suggests computer-aided methods for data processing and proposes alternative steps of analysis, it is clear that each of these contributions can only be one step in developing best practices. In the spirit of an open access policy, I would like to offer my data to the interdisciplinary community (see also the protocol of Study 2 in Appendix 4).

The working group 6 “Speech and other modalities” of the European Infrastructure Project CLARIN-D (www. http://clarin-d.de/) selected the data and metadata of Study 2 as their first motion-capture data set to be permanently accessible on an online platform, adapted to multimodal metadata standards (CMDI standards). It is hosted by the Bavarian Institute for Speech Signals (BAS) and can be retrieved under https://clarin.phonetik.uni-muenchen.de/BASRepository/Public/Corpora/NM-MoCap-Corpus/NM-MoCap-Corpus.1.php.

The data includes multi-perspective video data with synchronised motion-capture data and processed variables (see Section 9.9) of eighteen participants. Further, the Elan Plugin for Automatic Annotation (EPAA), the Motion-capture Java script (MoJa) and the Data Extractor (DEx) are part of the repository. For researchers around the world, the repository is freely available. For more information, consult the author (g@juliushassemer.de).

---

78 I, again, would like to thank Matthias Priesters, for his work on the metadata, on parts of the data and on the repository in its current form.
### 3. Background information participants (excerpt)

<table>
<thead>
<tr>
<th>Sex</th>
<th>Age</th>
<th>Occupation</th>
<th>Level</th>
<th>Field</th>
</tr>
</thead>
<tbody>
<tr>
<td>w</td>
<td>19</td>
<td>Student</td>
<td>BA</td>
<td>Machine building</td>
</tr>
<tr>
<td>m</td>
<td>20</td>
<td>Unemployed</td>
<td></td>
<td></td>
</tr>
<tr>
<td>w</td>
<td>21</td>
<td>Student</td>
<td>BA</td>
<td>Architecture</td>
</tr>
<tr>
<td>m</td>
<td>23</td>
<td>Student</td>
<td>MA</td>
<td>Informatics</td>
</tr>
<tr>
<td>m</td>
<td>23</td>
<td>Working on Staatsexamen</td>
<td></td>
<td>Medicine</td>
</tr>
<tr>
<td>w</td>
<td>24</td>
<td>Student</td>
<td>BA MA</td>
<td>History and Sociology</td>
</tr>
<tr>
<td>w</td>
<td>24</td>
<td>Student</td>
<td>MA</td>
<td>Architecture</td>
</tr>
<tr>
<td>m</td>
<td>26</td>
<td>Student</td>
<td>BA</td>
<td>Informatics</td>
</tr>
<tr>
<td>m</td>
<td>27</td>
<td>Security Engineer</td>
<td></td>
<td></td>
</tr>
<tr>
<td>m</td>
<td>27</td>
<td>Student</td>
<td>MA</td>
<td>Communication Science</td>
</tr>
<tr>
<td>w</td>
<td>29</td>
<td>Student</td>
<td>PhD</td>
<td>NA</td>
</tr>
<tr>
<td>w</td>
<td>30</td>
<td>Manager</td>
<td></td>
<td></td>
</tr>
<tr>
<td>m</td>
<td>30</td>
<td>Student</td>
<td>PhD</td>
<td>Informatics and Psychology</td>
</tr>
<tr>
<td>m</td>
<td>33</td>
<td>PostDoc Researcher</td>
<td></td>
<td>Linguistics</td>
</tr>
<tr>
<td>m</td>
<td>36</td>
<td>Student</td>
<td>MA</td>
<td>German and History</td>
</tr>
<tr>
<td>w</td>
<td>37</td>
<td>PostDoc Researcher</td>
<td></td>
<td>Architecture and Art</td>
</tr>
<tr>
<td>m</td>
<td>43</td>
<td>Professor</td>
<td></td>
<td>Linguistics</td>
</tr>
<tr>
<td>w</td>
<td>44</td>
<td>Secretary</td>
<td></td>
<td></td>
</tr>
<tr>
<td>m</td>
<td>48</td>
<td>Mine Engineer</td>
<td></td>
<td></td>
</tr>
<tr>
<td>m</td>
<td>51</td>
<td>Violin maker</td>
<td></td>
<td></td>
</tr>
<tr>
<td>w</td>
<td>55</td>
<td>Secretary</td>
<td></td>
<td></td>
</tr>
<tr>
<td>w</td>
<td>59</td>
<td>Marketing Manager</td>
<td></td>
<td></td>
</tr>
<tr>
<td>w</td>
<td>64</td>
<td>Pensioner</td>
<td></td>
<td>Tech. Med. Assistent</td>
</tr>
</tbody>
</table>

To not disclose personal data, only limited participant data are displayed.
### 4. Protocol

<table>
<thead>
<tr>
<th>Task</th>
<th>What on?</th>
<th>Time</th>
<th>Trial name</th>
<th>Cue for new trial</th>
<th>Text</th>
<th>Comments</th>
</tr>
</thead>
<tbody>
<tr>
<td>Preparation</td>
<td>MoCap operator: Prepares the study according to the checklist. During the study 3D-view and cameraviews on the screen!</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Consent form, handedness invent.</td>
<td>Interviewer welcomes participant, offers water. Explains and lets the participant read the consent form. Lets the participant fill out the handedness inventory. MoCap Operator puts on the Minimal Marker Set.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>How do you feel?</td>
<td>warming up</td>
<td>01:00</td>
<td>001</td>
<td>Participant enters volume</td>
<td>Wie fühlen Sie sich?</td>
<td>MoCapOp.: always presses ctrl+r if markers are weird</td>
</tr>
<tr>
<td>route description from big green door to here</td>
<td>warming up/ directions</td>
<td>02:00</td>
<td>002</td>
<td>&quot;Bitte beschreiben Sie mir ganz genau...&quot;</td>
<td>Haben Sie gut hierher gefunden? Bitte beschreiben Sie mir ganz genau, wie Sie unten von der schweren grünen Eingangstür bis hier in unsere Räume gekommen sind. So detailliert wie möglich.</td>
<td></td>
</tr>
<tr>
<td>1) Objects description</td>
<td>random object 1</td>
<td>00:30</td>
<td>011</td>
<td>&quot;Wie sah der Gegenstand aus?&quot;</td>
<td>Setzen Sie sich bitte aufrecht und entspannt hin und nehmen Sie diesen Gegenstand mit der flachen Seite zwischen Ihre Hände. Ich werde Ihnen jetzt kurz einen Gegenstand geben. Dann werde ich ihn wieder weglegen und Sie bitten, mir den Gegenstand zu beschreiben. Es gibt dabei eine Schwierigkeit: Bitte kommunizieren Sie so mit mir, als wäre ich ein Außerirdischer, der keine Gegenstände, Maßeinheiten und Formen kennt (Example &quot;chair&quot;). Sie dürfen aber Ihre Hände einsetzen. An Farbe, Geruch, und Gewicht bin ich nicht interessiert. Für jede Beschreibung haben Sie 20 Sek. Zeit. Ich frage Sie vorher ob Sie bereit sind. Auf meine Frage geht es dann los. <strong>Hand over the first object... take away.</strong> Wie sah der Gegenstand aus?</td>
<td></td>
</tr>
<tr>
<td>Calibration task</td>
<td>-</td>
<td>02:00</td>
<td>021</td>
<td>&quot;Und bitte!&quot;</td>
<td>Setzen Sie sich bitte aufrecht und entspannt hin und nehmen Sie diesen Gegenstand mit der flachen Seite zwischen Ihre Hände. <strong>Demonstrate</strong> - Die Unterarme waagerecht, Handfläche aufliegend aber entspannt. <strong>Hand over object</strong> - Und bitte!</td>
<td></td>
</tr>
<tr>
<td>Break:</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Participant takes off the markers. Interviewer sets up the screen, downloads the already uploaded clips and opens first clip in Elan.</td>
<td></td>
</tr>
<tr>
<td>2) Gesture observation. Show entire clip of first object. Let p. guess what object that is about.</td>
<td>Whole clip on random object 1</td>
<td>00:20</td>
<td>101</td>
<td>Um welchen Gegenstand geht es in diesem Video?</td>
<td>Ich werde Ihnen jetzt vorspielen was Sie mit den Händen während Ihrer Beschreibung gemacht haben. Wir schauen uns den Clip einmal komplett an und ich bitte Sie zu raten, um welchen Gegenstand es geht. Dann gehen wir alles Stück für Stück durch und ich frage Sie was die Hände hier tun. Hier kommt die Beschreibung des ersten Gegenstand s. <strong>Show whole video.</strong> Um welchen Gegenstand geht es in dies em Video? Dieses Video ist nicht das zuerst Aufgenommene. Die Reihenfolge ist zufällig.</td>
<td></td>
</tr>
<tr>
<td>First Gesture: 1. Does what that have to do with the object. If so: 2. What do the hands do? 3. Size/shape questions</td>
<td>random object 1, gesture 1</td>
<td>00:30</td>
<td>102</td>
<td>&quot;Hat das was die Hände machen etwas mit dem Gegenstand zu tun?&quot;</td>
<td>1. Hat das was die Hände machen etwas mit dem Gegenstand zu tun? (2. Was tun die Hände hier?) 3. Kann man dem was die Hände tun eine Größe/Bildung entnehmen? Kreisrand, auch wo wie eine gebogene Linie - auch wenn sie sich zum Kreis schließt? Zylinder rund wie ein hohler Zylinder oder ein Teil davon? Kugelrand wie eine Kugel oder ein Teil davon?</td>
<td></td>
</tr>
<tr>
<td>&quot;Are there any other landmark features?&quot;</td>
<td>gesture 2</td>
<td>00:30</td>
<td>103</td>
<td>&quot;Hat das etwas mit...&quot;</td>
<td>Wenn Sie finden, dass die Sequenz noch zum Vorigen gehört, können Sie das gerne sagen...</td>
<td></td>
</tr>
<tr>
<td>Now the second object...</td>
<td>random object 2</td>
<td>00:20</td>
<td>111</td>
<td></td>
<td>Kommen wir nun zum 2. Gegenstand...</td>
<td></td>
</tr>
</tbody>
</table>

---

268
5. Objects

<table>
<thead>
<tr>
<th>Object Name</th>
<th>Material</th>
<th>Proportions (cm)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Egg</td>
<td>Plastic</td>
<td>7.5/5</td>
</tr>
<tr>
<td>Drop</td>
<td>Plastic</td>
<td>18/8.5</td>
</tr>
<tr>
<td>CutOffPlate</td>
<td>Plastic</td>
<td>5/1.5</td>
</tr>
<tr>
<td>OvalPlate</td>
<td>Wood</td>
<td>14/10/0.3</td>
</tr>
<tr>
<td>Object</td>
<td>Material</td>
<td>Dimensions</td>
</tr>
<tr>
<td>------------------</td>
<td>--------------</td>
<td>------------</td>
</tr>
<tr>
<td>Round Plate</td>
<td>Wood</td>
<td>4/0.5</td>
</tr>
<tr>
<td>Hollow Cylinder</td>
<td>Hard cardboard</td>
<td>8/5/0.3</td>
</tr>
<tr>
<td>Big Sphere</td>
<td>Metal</td>
<td>8</td>
</tr>
<tr>
<td>Small Sphere</td>
<td>Wood</td>
<td>2</td>
</tr>
<tr>
<td>Torus</td>
<td>Wood</td>
<td>5.5/1</td>
</tr>
</tbody>
</table>
6. Settings for Vicon Nexus automatic post-processing
7. Additional tokens of Shape Gestures adding to Figure 10.07

Measure Gestures (Study 2)

Shape Gestures (Study 2)