Producing and perceiving gestures conveying height or shape

Winter, Bodo; Hassemer, Julius

DOI:
10.1075/gest.15.3.07has

Document Version
Peer reviewed version

Citation for published version (Harvard):
Winter, B & Hassemer, J 2016, 'Producing and perceiving gestures conveying height or shape', *Gesture*, vol. 15, no. 3, pp. 404. https://doi.org/10.1075/gest.15.3.07has

Producing and perceiving gestures conveying height or shape

Julius Hassemer[a] & Bodo Winter[b]

[a] Independent researcher

[b] University of California, Merced, Cognitive and Information Sciences

Abstract

In this paper, we analyze single-handed hold gestures that convey the height or the shape of an object ('height gestures' and 'shape gestures'). The analyses include (1) differentiating which parts of the hands are profiled in each gesture, (2) considering whether these parts are occluded by other fingers, and (3) deriving predictions from (1) and (2) with regard to what hand shape characteristics favor height or shape gestures. In a production experiment, we asked participants to indicate either the height of a small ball or its shape, using only the index finger and the thumb. As predicted, the middle, ring and little finger were more curled in for height gestures. For shape gestures, those three fingers were more raised. In two perception experiments, participants viewed virtual hands from a computer-generated continuum of hand-shape stimuli. We found the same pattern of that if the middle, ring and little finger were more curled in, participants were more likely to interpret a gesture as indicating height. Our results pose new questions with regard to the relation between the physical form of a gesture and the inferred imaginary forms that are involved in deriving meaning from a gesture.

1. Introduction

When people talk about physical objects, they frequently use gestures to characterize them. For example, a speaker might say "it was a pebble of this size" and move the index finger close to the thumb, as if demarcating a small distance between the fingers. Or, a speaker might say "the pebble had this round shape" while curving the fingers in a C-shape, as if enclosing a round object.

It is, of course, undisputed that there are many ways through which gestures can convey such spatial properties as shape or size (e.g., Andrén, 2010; Beattie & Shovelton, 2006; Calbris, 2003; Hassemer, 2016; Holler, Shovelton, & Beattie, 2009; Müller, 1998; Sowa, 2006; Streeck, 2009). This paper adds to this body of research by focusing on how gesturers encode spatial information through specific hand shapes, and how observers decode this information. We focus on the contrast between height and shape: What aspects of the hand shape matter in conveying height? What aspects matter in conveying shape?

Gesture observers do not solely perceive the three-dimensional shape and movement of the hand. Instead, they seem to immediately interpret physical hand movements as indicating something that goes beyond the mere physical articulation (Calbris, 1990; Cienki, 2005; Mittelberg, 2010; Müller, 1998). Or, in other words, "The gesture presents an image of the invisible" (McNeill, 1992, p. 14). For example, think about moving the index finger along an imaginary line in the air: Instead of merely perceiving an index finger at different spatial positions, one tends to perceive a "geometric path described by the finger-*tip*" (James, 1890, p. 190, italics in original). Gestures have the ability to create such "virtual objects, abstracted from motions of the hands" (Streeck, 2008, p. 291); they can, as in the example of drawing a line in the air, leave "imaginary traces in gesture space" (Mittelberg, 2010, p. 367). Hassemer (2016) and Hassemer, Joue, Willmes, and Mittelberg (2011) distinguish between 'gesture form' and 'physical form'. Gesture form includes imaginary (but crucial) forms that are abstracted from the concrete physical form of the gesturing articulator, for example, an imaginary line inferred from continuous movement of the finger.

Here, we apply the theoretical distinction of gesture form and physical form to gestures that are performed by the extended index finger and thumb with their tactile surfaces facing each other, similar to "gestures of precision grip" (Kendon, 2004, Ch. 12; see also Streeck, 2009, Ch. 3). Gestures from the 'precision grip' family can have many different meanings. For example, we can use such gestures to express the importance or precision of what is being said (Kendon, 2004; Lempert, 2011), or to highlight a small quantity or size (Winter, Perlman, & Matlock, 2013). In this paper, we look at how the index finger and the thumb partially enclose an imaginary object ("external metonymy", Mittelberg & Waugh, 2009) and may indicate either the height or the shape of an object (Hassemer, 2016; Sowa, 2006).

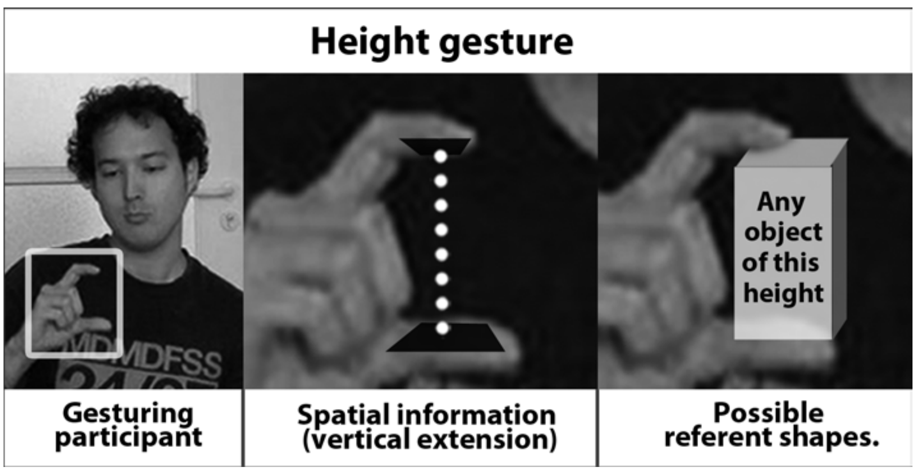
In the following section, we discuss the distinction between physical form and gesture form within the framework of Gesture Form Analysis (§2), focusing on how these two concepts relate to each other in index-finger-and-thumb gestures that express either height or shape. We then report on the results of a production experiment, for which we photographed prompted gestures. This data shows that people produce a systematic contrast between height and shape gestures in line with predictions from Gesture Form Analysis (§3). Following this, we report on two perception experiments, a binary forced-choice task (§4.1) and a free drawing task (§4.2). In both of these tasks, we manipulated key aspects of the hand configuration using virtual hands. The results show that the position of fingers that are not in the foreground of attention (the middle, ring and little finger) determine whether a gesture conveys height or shape.

2. Background

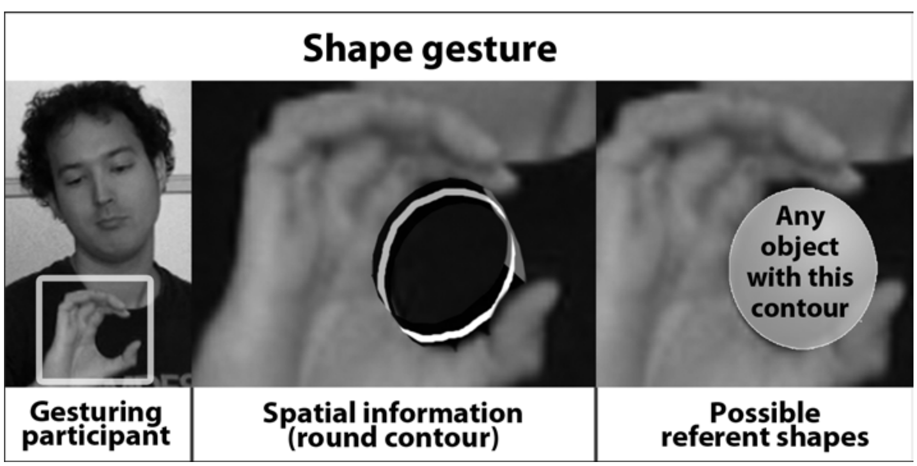
2.1. Physical form versus gesture form

We use the framework of Gesture Form Analysis (Hassemer, 2016; Hassemer et al., 2011) to understand how a gesturer or a gesture observer goes from merely seeing a hand shape to deriving abstracted height or shape information. This framework is the attempt to structure gesture conceptualization by analyzing its necessary spatial operations. A basic tenet of Gesture Form Analysis is that gestures are not only interpreted via the physical shape of the articulator (see, e.g., Stokoe, 2005/1960; Bressem, 2013), but that the physical form is also systematically processed to arrive at a more abstract form (Sowa, 2006; Mittelberg & Waugh, 2009). For drawing a line in the air, the movement of the hand from left to right with an extended index finger is the gesture's physical form. But this particular movement and configuration leads to gesture form, which consists of the physical form being processed in various steps. First, rather than merely perceiving the whole hand or the entire body, the index finger is profiled (i.e., attention is allocated to it). The focus is particularly on the point at the tip of the finger. Then, this moving point – by perceiving its motion as "atemporal motion" (Mandel, 1977) or "trace leaving" (Hassemer, 2016) – creates a linear form, an imaginary line. This line may have further meaning in the discourse context, for example, it may indicate the (metaphorical) ups and downs of a romantic relationship (Müller, 2008, p. 235), or it may enclose a two-dimensional form such as the contour of a round picture frame (Müller, 2014, p. 1690). Gesture form thus spans a succession of spatial operations that, in consequence, create an imaginary line in the air: from a hand with extended index finger, to the point at its tip, to the trace the point leaves (for a complete analysis, see Hassemer, 2016). The proposal of a succession of spatial operations is supported by Talmy's work on the conceptualization of pointing gestures as a "fictive chain" from the gesturer to the target (Talmy, forthcoming).

The distinction between physical form and gesture form also applies to gestures that are not moved during the expressive phase, so-called "independent hold" gestures (Kita, van Gijn, & van der Hulst, 1998). Two particular independent hold gestures will be of interest for this paper: First, a 'measure gesture' or 'height gesture' which indicates a referent's extension along one axis (in this paper: vertical extension). Second, a ‘shape gesture’, which indicates a referent’s outer contour via a C-shape (Hassemer, 2016). Figure 1 contrasts these two gestures and highlights the aspect of a Gesture Form Analysis that are relevant for this paper: Different surface areas are profiled for each of the two gestures (see Sowa, 2006, p. 199). In a height gesture interpretation, only the two opposing surfaces of the index finger and thumb pad are profiled, with the resulting form being specified by the vertical distance between the two surfaces. Shape gestures, on the other hand, involve profiling the complete inner surface of the index finger, the thumb, and the side of the hand between these fingers. This single inner surface in the shape of a “C” encloses the resulting form, thus conveying its contour.



a)



b)

Figure 1: (a) Height gesture and (b) shape gesture. The middle, ring and little finger are curled in for the height gesture.

In contrast to the index-finger-line example belonging to the gesture type "drawing" (Hassemer, 2016), the height and shape gestures are classified as "holding" gestures (cf. "modes of representation" of "acting, molding, drawing, representing", Müller, 2014, p. 1687). Holding, in Hassemer (2016), is defined by the hand being in static "surface contact" with an imagined object, as if grasping the object. Grasping an object involves at least partially enclosing the object's shape. The gestures in Figure 1 show two of many ways to partially enclose an imagined referent: The height gesture touches the referent at two extremes; the shape gesture encloses the referent more fully, spanning large parts of its contour.[1]

The imaginary line example discussed above involved creating a one-dimensional form that had to be inferred particularly from the movement of the hand. The holding gestures shown in Figure 1, on the other hand, are static, but nevertheless convey a referent's shape. If movement was added to the hand shapes shown in Figure 1, a referent's shape and motion could be signaled simultaneously, all in one gesture. This is not possible for a drawing gesture (whose movement is already 'occupied' with characterizing the shape of the line). Thus, the parameter "motion" is optionally available in height and shape gestures, but these gestures do not require motion to succeed in signaling height or shape.

Whereas height gestures specify a referent in its vertical dimension only, shape gestures specify the referent in two dimensions. Shape gestures can refer to various round objects including a disc, a ball or a cylinder. Both in drawing and in holding, the physical form of the articulator is somewhat of secondary importance and mainly serves the purpose of indicating gesture form; thus disclosing specific spatial properties of a referent.

It should be emphasized that height and shape gestures constrain the reference object in different ways: The shape gesture can indicate objects of different sizes. Since the gesture focuses on shape, rather

[1] Both height and shape gestures can be performed in front of the body with fingers 3-5 facing away from the body (Figure 1) or with an extended arm besides the body (with the index and thumb facing away from the body; see Figure 2).

than on height, one could use it when describing a small ball (fitting into the hand) or a very large ball (not fitting into the hand). On the other hand, the height gesture specifies the size of the intended referent; but it is neutral to the shape of the referent. The referent could be rectangular, but it could also be of any other form, for example, in the shape of a key or a small figurine. Moreover, it could be any of the round forms acceptable for the shape gesture. This difference in flexibility regarding shape and size is highlighted in Figure 1 in the right column, with height gestures annotated as indicating "any object of this height", thus under-specifying shape; and shape gestures annotated as indicating "any object with this contour", thus potentially under-specifying size.

For the experiments reported below, height and shape gestures were selected because they are similar in their physical form (as discussed in Sowa, 2006), but clearly diverge in the spatial information they convey. This allows studying what aspects of the physical form give rise to differences in gesture form, i.e., what determines a height or shape interpretation in the gesturer's or observer's mind. Moreover, we chose to study these gestures because they occur frequently. They are often used to describe hand-sized objects (see §2), and they are also part of size and shape classifiers in many sign languages (e.g., Zwitserlood, 2012, p. 16). However, these gestures are clearly not *only* height or shape gestures, that is, they can be used in other contexts as well. Gestures are highly multifunctional (e.g., Calbris, 2011); for example, precision grips similar to the gestures studied here can co-signal pragmatic meanings (Kendon, 2004, Ch. 12; Lempert, 2011) together with spatial and metaphorical meanings (Winter et al., 2013). Our experiments do not address the multifunctionality of gesture, and they do not contradict it either; they merely focus on those specific contexts where the spatial properties of objects are salient.

2.2. Predictions about hand shapes

When it comes to the physical form of the hand, two main hand-shape variables are predicted to be relevant for distinguishing between height and shape gestures. The first and perhaps most important variable is what we will refer to as PINKIE CURL. This variable refers to the angle of the middle, ring and little finger (henceforth 'fingers 3-5') in relation to the palm. Height gestures are predicted to have fingers 3-5 more curled in, as displayed in Figure 1a. Curling in the fingers 3-5 interrupts the C-shape because the fingers intrude into the area formed by the C. This focuses attention not on the C-shape, but on the profiled index finger and thumb pads. For shape gestures, on the other hand, fingers 3-5 are predicted to be raised, not occluding the C-shape. The second physical variable is the degree of curvature of the index finger, INDEX CURVE. As illustrated by the different hand shapes in Figure 1, the index finger is expected to be less curved in height gestures[2] and more curved in shape gestures—at least for round objects. Thus, we predict for each variable:

(1) PINKIE CURL: Fingers 3-5 more curled in for height than for shape gestures

(2) INDEX CURVE: Index finger more curved for shape than for height gestures

Both of these predictions rely on the existence of gesture form (the height or the contour conveyed to an observer) as a derivation of physical form (the gesturer's hand shape).

These predictions were already borne out in a production study presented in Hassemer (2016), where 23 German-speaking participants were recorded by optical motion-capture cameras while describing nine hand-sized objects.[3] Measure gestures and shape gestures were coded by three independent raters (with high inter-rater reliability, Fleiss' κ=0,80). In the entire set of 207 trials, there

[2] The thumb tends to curve with the index finger. Since different degrees of curvature are more clearly discernible in the index finger than in the thumb, we will focus on the index finger in the following. However, it should be kept in mind that the thumb curves as well.
[3] Video and 3D data, codings, tools and metadata (CMDI standards) of the motion-capture study are available at the Bavarian Institute for Speech Signals (BAS), as part of the curation project "Editing and Integration of multimodal resources in CLARIN-D" under the link hdl.handle.net/11858/00-1779-0000-0006-BF00-E

was a total of 471 measure and shape gestures (including repetitions), indicating that these are indeed two frequent gesture types in describing objects. Tracking motion-capture markers on the hand allowed quantifying PINKIE CURL and INDEX CURVE precisely (see Hassemer, 2016 for details on methods), and both variables were significantly affected by the gesture type in the predicted direction. In particular, for PINKIE CURL, the mean angles of measure gestures (83°) and shape gestures (36°) differed by 47°, indicating that fingers 3-5 were much more curled in for the measure gestures.

A shortcoming of the motion-capture study is its unconstrained nature. There were gestures about different objects (e.g., small ball, large oval plate) and about different spatial aspects of the objects (e.g., height of the whole plate versus thickness of its material). Further, many of the gestures were complex. For example, they combined shape and height gestures with movements tracing over larger parts of the imaginary object. In the next section, we present results from a more controlled follow-up experiment with a single object, for which we prompted participants explicitly to either indicate the height or the shape with a single-handed gesture. This shows that the height/shape contrast is indeed a regular and systematic pattern produced by gesturers. In the analysis of the production experiment, we focus on the role of the variable PINKIE CURL. We return to the other variable, INDEX CURVE, in the perception experiments below.

3. Producing height and shape gestures: A photography task on prompted gestures

3.1. Stimulus and procedure

We asked 55 participants to indicate the height and the shape of an object. An interviewer (first author) asked German-speaking pedestrians in Berlin Friedrichshain and Lichtenberg whether they were available for a very short experiment, conducted on the spot. Participants were shown a Styrofoam ball of 6 centimeters in diameter (Figure 2a). There is no particular reason to prefer a ball-shaped object to a round plate or another round form, as done in Hassemer (2016); only a ring or another shape that can be instantiated by the solid form of the fingers themselves (in contrast to enclosing the shape) might be problematic.

The interviewer presented the ball on a flat hand before hiding it in a bag. After presentation, the participants were either asked to show the height of the object by using the index finger and thumb of one hand; or they were asked to show its shape with the same fingers. Each participant was asked both the height and the shape question (the question order was randomized across participants). After debriefing the participant, the next pedestrian outside of hearing distance was asked to participate.

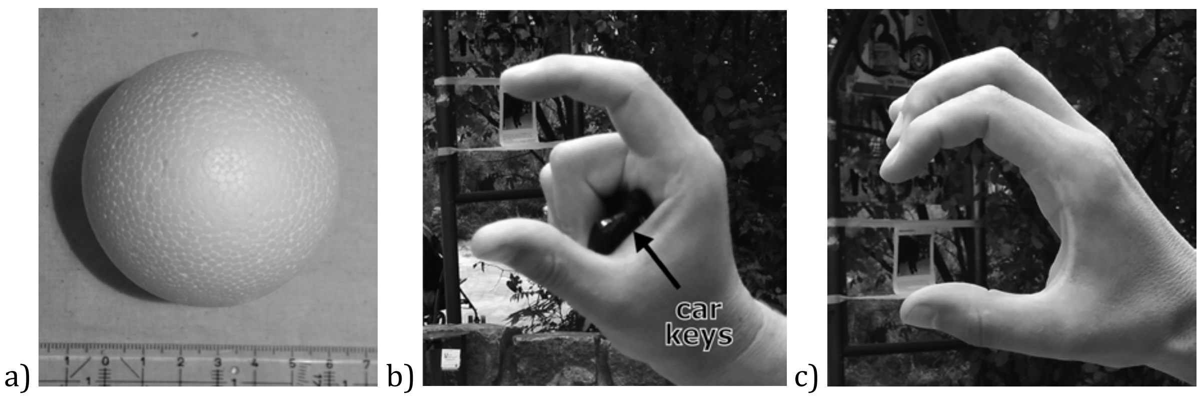


Figure 2: (a) Stimulus object for the production experiment. (b) A participant performing a height gesture with his right hand in response to being asked to indicate the object's height. (c) The same participant performing a shape gesture with his right hand, with fingers 3-5 less curled in/more raised. Notice that in (b), the speaker holds his car keys between fingers 3-5, which he puts into his left hand before producing the shape gesture in (c). See results (§3.3) for a discussion of this specific example.

3.2. Analysis

Focusing on still images disregards preceding and following movements, which may create more complex forms by 'leaving a trace' (§2), and the hand configuration might vary slightly during the holding phase. However, since the primary variable of interest is a rather static variable characterizing hand shape

(PINKIE CURL), we regard the use of still images as acceptable in this case. In real discourse, of course, the hand configuration is embedded in a continuous stream of gesturing.

Since the motion-capture study already indicated PINKIE CURL to be a *metric* variable distinguishing the two gesture categories (§2.2), we simply coded whether fingers 3-5 were curled in more for height or more for shape gestures. Two coders looked at both photographs from each participant (110 pictures in total), coding for whether fingers 3-5 are curled in more for the first or the second picture. The coders did not know the order of questioning for a given pair of photos. We used the R package "irr" (Gamer, Lemon, & Singh, 2012) to calculate inter-rater reliability, which was "almost perfect" (Cohen's $\kappa = 0.926$; Landis & Koch, 1977).

3.3. Results

All statistical analyses were conducted with R (R Core Team, 2014). In line with standards for reproducible research in data science (Gentleman & Lang, 2007; Mesirov, 2010; Peng, 2011), the scripts and all data (including stimulus pictures and codings) are made available with this publication and can be retrieved online[4].

As predicted, the hand configurations with more PINKIE CURL were responses to the height question; hand configurations with less PINKIE CURL were responses to the shape question (simple Chi-square test: $\chi^2(1)=24.8$, $p<0.0001$). A total of 85% of the participants (N=47) produced gestures that matched our hypothesis. Conversely, only 15% of participants (N=8) used hand shapes with more PINKIE CURL for shape gestures. Thus, people systematically curled in fingers 3-5 more when asked to indicate shape. When asked to indicate height, fingers 3-5 were relatively more extended.

There is a noteworthy anecdote from one specific participant who was holding his car keys in his right hand when he was approached by the experimenter. The height question was asked first for this

[4] http://www.github.com/bodowinter/height_shape_gestures/

participant, and the participant held his right hand as shown in Figure 2b, with the car keys still in this hand. The car keys were enclosed by fingers 3-5. Interestingly, when asked the shape question immediately afterwards, the participant put the keys into his left (non-gesturing) hand. After this, he moved his right hand up again to produce the shape gesture shown in Figure 2c. Although anecdotal, this example hints at the importance of the C-shaped contour in shape gestures: Apparently, the participant did think of the car keys as interfering with the shape gesture, but not with the height gesture. Thus, the participant moved the car keys out of the gesturing hand precisely for the shape gesture that involves profiling the entire inner surface of the C-shape.

Having established that prompting for height or shape does indeed produce two recurring gesture types distinguished by PINKIE CURL, we are now going to investigate how specific hand-shape variables (PINKIE CURL and INDEX CURVE) affect the *perception* of height and shape gestures.

4. Perceiving height and shape gestures

4.1. Binary forced-choice task

4.1.1. Stimuli & procedure

We used the 3DS MAX modeling and rendering software to create a virtual hand (courtesy of the 3D designer Philipp Krecklow). Using this hand model, we manipulated the PINKIE CURL variable along a nine-step continuum, with increasing curl of fingers 3-5. Orthogonal to this, we manipulated the INDEX CURVE variable along a six-step continuum, with increasing curvature of the index finger. This resulted in 9 * 6 = 54 total stimuli, displayed in Figure 3. The PINKIE CURL continuum had more steps than the INDEX CURVE continuum because the PINKIE CURL of fingers 3-5 represents a larger movement, with larger visual differences and more degrees of freedom than the INDEX CURVE continuum.

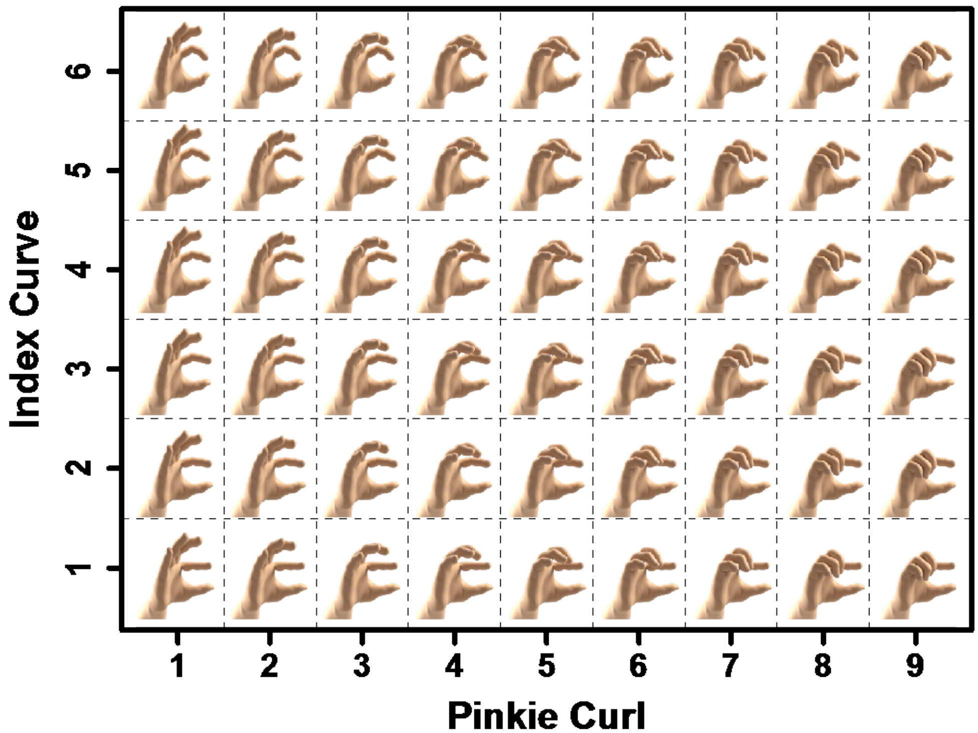


Figure 3: Hand-shape matrix showing all combinations of PINKIE CURL and INDEX CURVE in our perception experiments

A total of 309 participants of English speakers residing in the United States and Canada participated for a small reimbursement (20 US cents) on Amazon Mechanical Turk (http://www.mturk.com). This platform is known to be a valid tool for collecting behavioral data (Bohannon, 2011) and performing linguistic experiments (Sprouse, 2011). Each participant was given the following instructions:

> "On the next screen, you will be shown a picture of a hand for a few seconds. The gesture you will see characterizes an object. Please keep in mind what the hand looked like."

Before showing the image, we presented a counter (counting down from 3 to 1 in three seconds) to assure that all participants focused on the screen at approximately the same time. Then, one of the virtual hands was presented on the screen for three seconds. Following this, participants were asked:

> "The gesture you just saw characterized an object. What exactly do you think the gesture was about?
>
> (a) It showed the shape of an object.
>
> (b) It showed the height of an object."

The order of the (a) and (b) response options was randomized (RESPONSE ORDER). Once they provided their response, participants were asked how confident they felt about their choice on a scale from 1 (not at all confident) to 9 (very confident). Participants responded that they were relatively confident with their choices (7.13, significantly different from the midpoint 5; $t(308) = 25.58$, $p<0.0001$). Finally, we asked whether participants were able to view the picture (in case they might not have looked at the screen within the three-seconds period); all participants selected "yes".

4.1.2. Analysis

Since the response is a binary choice and every participant only contributed one data point (single-item between-subjects design), we used logistic regression to model the relationship between the dependent variable CHOICE (height versus shape response) and the independent variables INDEX CURVE and PINKIE CURL (the two variables were entered as continuous predictors, centered). We also entered two additional quadratic effects for INDEX CURVE and PINKIE CURL. This allows for the possibility that intermediate and extreme values of the continuum differ, for example, the results may follow a U-shaped pattern. We further entered RESPONSE ORDER (height first versus shape first) as categorical predictor (deviation coded). Finally, we entered an interaction of the linear terms of INDEX CURVE and PINKIE CURL. This is motivated in analogy to categorical perception experiments, where oftentimes, different cues for a linguistic phenomenon may interact, as in perceptual trading relations (Repp, 1982; Parker, Diehl, & Kluender, 1986).

An initial model revealed that neither the INDEX CURVE * PINKIE CURL interaction ($\chi^2(1)=0.13$, $p=0.72$), nor the control variable RESPONSE ORDER ($\chi^2(1)=1.15$, $p=0.28$) were significant. Hence, we simplified the model by excluding these predictors. Simplifying the model did not change the significance of any of the effects reported below. The following p-values are based on likelihood ratio tests.

4.1.3. Results

Overall, there was a response bias, with more participants responding "shape" than "height." A total of 60% (N=184) of the participants indicated that the virtual hand showed the shape of an object. Only 40% (N=125) indicated that it showed the height of an object. Crucially, whether participants chose the shape or height response option was significantly affected by both INDEX CURVE and PINKIE CURL. Increasing INDEX CURVE lead to significantly more shape responses ($\chi^2(1)=41.5$, $p<0.0001$). For each increase in INDEX CURVE by one step along the virtual hand continuum, the odds of a shape response increased by 1.69 to 1 (log odd estimate of shape response: 0.53, SE=0.009). On the other hand, increasing PINKIE CURL lead to significantly more height responses ($\chi^2(1)=7.05$, $p<0.01$). For each increase in PINKIE CURL by one step along the virtual hand continuum, the odds of a height response increased by 1.15 to 1 (log odd estimate of height response: 0.14, SE=0.053).

The quadratic terms were also significant, for both INDEX CURVE ($\chi^2(1)=4.16$, $p<0.05$) and PINKIE CURL ($\chi^2(1)=20.31$, $p<0.0001$). This suggests that INDEX CURVE and PINKIE CURL affected the perception of spatial properties in a nonlinear way. Figure 4 plots the predictions of the logistic regression for both variables (marginal plots, ignoring other factors). As can be seen, the nonlinearity is particularly noteworthy for PINKIE CURL: As predicted by Gesture Form Analysis, when fingers 3-5 are progressively more curled in, people are progressively more likely to perceive the hand shape as indicating the height of an object. This can be seen by the fact that only steps 8 and 9 have significantly more height than shape responses, i.e., the 95% confidence intervals are above the 50% midline. In other words: Overall there is a bias to interpret the hand shape as a shape gesture, but this bias is counteracted by very high PINKIE CURL values. But, contrary to our predictions, very low values of PINKIE CURL (i.e., the fingers 3-5 are very much raised) also had a somewhat higher proportion of "height" responses. One potential explanation for this is that if fingers 3-5 are close to the index finger (medium PINKIE CURL values, as exhibited in Figure 1b), they might be seen as loosely mirroring the index finger curvature, thus reinforcing the round contour that is the core of a shape gesture interpretation. When fingers 3-5 are raised even more (very low PINKIE CURL), this effect is not present, and height gestures are at least equally likely.

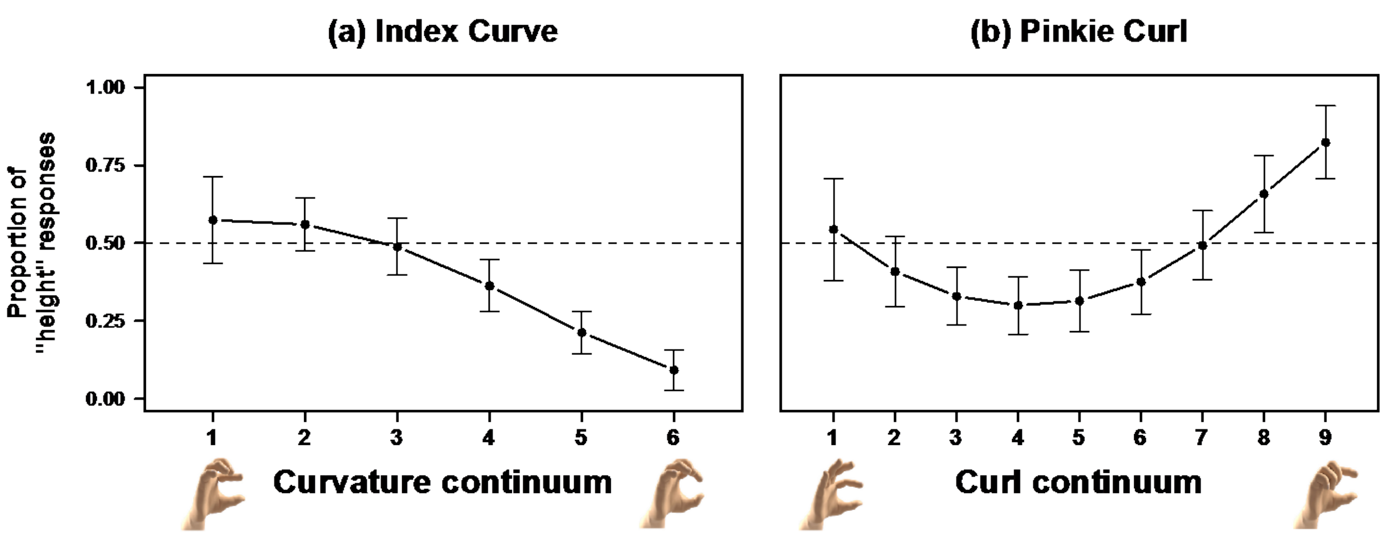


Figure 4: (a) Proportion of height responses for the INDEX CURVE continuum. The points indicate predictions of the logistic regression model with 95% confidence intervals. Higher values on the x-axis indicate more curved index fingers. The displayed hand shapes represent the end points of the INDEX

CURVE continuum (with a medium PINKIE CURL value of 5). (b) Proportion of height responses for the PINKIE CURL continuum (logistic regression fit). Higher values on the x-axis indicate fingers 3-5 being more curled in. The hand shapes represent the end points of the PINKIE CURL continuum (with an INDEX CURVE value of 3). In both images, the dashed line indicates a proportion of 0.5 (equal probability of responding height or shape). As can be seen, height responses were more likely than shape responses only for low INDEX CURVE, high PINKIE CURL and very low PINKIE CURL.

The forced binary-choice experiment clearly shows that INDEX CURVE and PINKIE CURL influenced participants' responses with respect to the height/shape perceptual contrast. However, due to the forced-choice nature of the task, we could have inadvertently biased our participants, since we imposed a binary distinction that emphasizes the contrast that we were interested in (height versus shape). In the following task, we asked participants to freely draw objects in response to the same stimuli. This means that this time, we do not prime the participants by mentioning the concepts of height or shape. Moreover, we allow for responses that are potentially intermediate, something that was not possible in the forced-choice task since there was no response option "height *and* shape". If we find evidence for the contrast between height and shape gestures even if participants are free to draw whatever they want, this could show that there is indeed a recurring distinction between these types even in the unprompted interpretation of hand shapes.

4.2. Free drawing task

4.2.1. Procedure & analysis

We printed black-and-white versions of the 54 hand pictures used in the previous experiment. A total of 168 drawings were collected, each one from a different participant (undergraduate students from the

University of California, Merced). Participants performed the task for course credit after they completed an unrelated experiment. On the top of the page, the instructions read:

> "This hand characterizes an object. Please draw onto this image what you think the object could look like."

A total of 23 data points (14%) were excluded because they contained drawings that cannot be used to address our research question. This was primarily either because participants saw the hand as the referent itself (Figure 5a), or because participants did not draw an object onto the hand at all (Figure 5b).

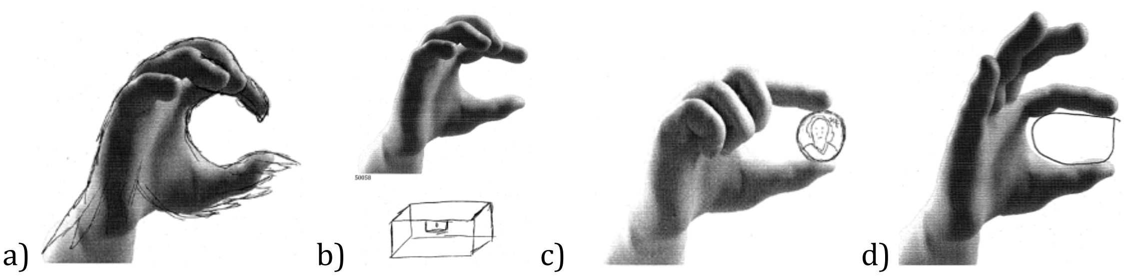


Figure 5: (a) A drawing with the hand itself being interpreted as representing the object (reason for exclusion). (b) A drawing with the object not drawn onto the hand (reason for exclusion). (c) A coin drawn in a position consistent with a height gesture interpretation. (d) An abstract, slightly oval shape drawn in a position consistent with a shape gesture interpretation.

Figure 5c and Figure 5d represent drawings that are consistent with height and shape gesture interpretations, respectively. We derived several variables, which were coded by three independent coders (Gurjot Bains, Neekole Arcoda, Clarizza Arucan). The variable POSITION was measured by the first author. The following variables were considered:

ROUNDEDNESS

Whether the object is round.

binary: round/non-round

SYMMETRY

Whether the object is drawn symmetrical to a vertical line between the center of the index finger and thumb pad.

binary: symmetrical/non-symmetrical

POSITION[5]

The position of the object's center on a horizontal axis.

continuous: millimeters

All binary variables were analyzed with logistic regression. POSITION was analyzed using simple linear regression. In all models, we entered INDEX CURVE and PINKIE CURL as continuous predictors (centered).

4.2.2. Results

4.2.2.1. The variable ROUNDEDNESS

PINKIE CURL did not significantly affect whether drawings were predominantly round or not round ($\chi^2(1)=5.85$, $p<0.05$). However, INDEX CURVE did affect ROUNDEDNESS ($\chi^2(1)=0.89$, $p=0.35$). For the smallest INDEX CURVE value, logistic regression indicates an estimate of only 32% round drawings. For the largest INDEX CURVE value, the logistic regression indicates an estimate of 61% round drawings. For every one

[5] For this variable only, we added data from 81 participants that were collected in a classroom. Every student was instructed not to look at the sheets from their neighbors.

step increase of the INDEX CURVE variable, the predicted odds of observing a round drawing increase by 1.27 to 1 (log odd estimate: 0.24, SE=0.1). This result is perhaps not surprising: Rounder index-finger-and-thumb configurations presumably primed participants to think of round objects. However, this serves to show that people did pay attention to the hand configuration and that their drawings did reliably relate to what was visible on the sheet. From this perspective, the influence of INDEX CURVE on ROUNDEDNESS acts as a manipulation check.

4.2.2.2. The variable SYMMETRY

INDEX CURVE did not significantly affect whether drawings were predominantly oriented along the index finger/thumb axis ($\chi^2(1)=0.4$, $p=0.52$), however, PINKIE CURL did affect SYMMETRY ($\chi^2(1)=7.5$, $p<0.01$). For the smallest PINKIE CURL value, the logistic regression indicates an estimate of only 30% of objects that are centered symmetrically around the vertical line between index finger and thumb. For the largest PINKIE CURL value, the estimate is 60%. For every one step increase in PINKIE CURL, the odds of observing a drawn object that is 'index symmetric' increase by 1.24 to 1 (log odd estimate: 0.21, SE=0.08). Thus, the results indicate that if fingers 3-5 are curled in, people are more likely to draw objects that are symmetrical around the index finger/thumb line. If fingers 3-5 are raised, people are less likely to draw objects centered on this line.

4.2.2.3. The variable POSITION

The variable POSITION documents the position on the horizontal axis of those drawings that have a clear center, i.e., drawings that are either symmetric to a point or symmetric to a vertical line. This position is tracked with reference to a fixed point on the picture of the virtual hand. Higher POSITION values mean that the participant drew the object more towards the right of the picture. The mean of POSITION is at 52mm (SD=10.7mm), which is close to the line between the index finger and thumb. Hartigan's dip test (implemented by the package "diptest", Maechler, 2015) indicates no significant deviation from unimodality (D=0.019; p=0.97), showing that the drawings are indeed distributed around a POSITION value of 52mm. Thus, despite a free choice when it comes to placing the object onto the sheet, participants tend to align their drawing with the vertical line between index and thumb pad.

INDEX CURVE did not affect the variable POSITION (F(1,162)=1.97, p=16). However, PINKIE CURL did have a significant effect on POSITION (F(1,162)=15.64, p<0.001). With each one-step increase along the PINKIE CURL continuum, the center of the object was on average 1.9mm more toward the right (SE=0.48). This confirms a central assumption of this paper: High PINKIE CURL values lead to the profiling of articulators farther to the right. For height gestures, not the entire C-shape in the center of the picture is profiled; instead, the finger pads on the right of the picture are profiled. Conversely, if the fingers 3-5 are extended, the C-shape is less occluded and more people draw images in the center of the space enclosed by index finger and thumb.

5. Discussion

The results confirm several predictions derived from the Gesture Form Analysis presented in §2: Curling in fingers 3-5 occludes part of the imaginary forms that are involved in interpreting a shape gesture, the C-shape inside the index finger and the thumb. The observer's attention being on the whole C-shape is thus unlikely, favoring the space between the index finger and the thumb pad as the focus of attention. In such a height gesture interpretation, only the finger pads are profiled. Contrastingly, hand shapes with non-occluded C-shapes and fingers 3-5 raised are perceived to indicate shape rather than height. More generally, the difference between height gestures and shape gestures was shown to be a meaningful contrast, one which influences how participants indicate an object's height or shape when prompted (§3), as well as a contrast that they reliably make use of when they interpret different stimulus hand shapes, both with a restricted, binary response option, and with free, continuous response possibilities (§4).

The evidence supports that a hand shape similar to so-called 'precision grip gestures' (but with the index finger and thumb typically not touching each other) can be differentiated into two distinct gesture types: height and shape gestures. In our experiments, participants reliably distinguished these two gestures, in production as well as in perception.

The detailed analysis of two gestures that are similar in their physical form but different in gesture form, that is their spatial interpretation, raises questions about the perception of gesture in contrast to perception in other domains, for example in speech. The forced-choice perception experiment in §4.1 comes closest to studies done on categorical speech perception (e.g., Liberman, Harris, Hoffman, & Griffith, 1957). In these studies, stimuli that vary along a continuum are presented to participants, who then have to report whether they perceive one category or another. Take, for example, the distinction between voiced and voiceless plosives, as in "bear" versus "pear". Continuous variation along the voicing continuum is perceived in a strikingly discontinuous fashion, with almost everybody perceiving a clear and unambiguous "bear" or "pear" when a certain threshold is passed. In our data, there is no such cut-off, and the response curves do not look nearly as sigmoidal as in categorical speech perception studies. At least with regard to the selected hand-shape variables, gesture appears to be more graded and analog than speech perception, despite persistent semi-categorical regularities in gesture production (§3). Our results also contrast to sign language perception as shown by studies on classifiers in American Sign Language, which also appear to be more categorical (Emmorey & Herzig, 2003) than the contrast investigated in this paper.

Nevertheless, the absence of a clear-cut categoricity in our data does not provide evidence for a general lack of categorical perception in gesture. In particular, the non-linear nature of the effect observed in the binary forced-choice task (§4.1) suggests that our experiment may implicitly test other perceptual factors. In this experiment, shape gesture interpretations peaked at a middle position of the PINKIE CURL continuum, something that we did not originally anticipate and that could be due to fingers 3-

5 visually reinforcing the curvature of the index finger in this configuration. For an experiment that is truly comparable with those done in categorical speech perception, all such additional perceptual factors would have to be controlled for.

Regardless of the issue of categoricity, the results of this paper flesh out the contrast between physical form and gesture form discussed in §2. With gesture, the physical form is necessarily always the basis for any interpretation. But particular aspects of this physical form cue a specific gesture form interpretation for example in that they favor profiling different parts of the hand (see also different "gestural cues" in Talmy, forthcoming). In our case, the position of fingers 3-5 was the primary variable that shifted the focus away from the inner C-shaped contour. When this happens, the focus shifts to the index finger and thumb pads (§4.1 and §4.2). These results show how different fingers interact with each other to create gesture form: In both height and shape gestures, parts of the index finger and thumb are profiled. However, it is the fingers *other than the index finger and thumb*, namely fingers 3-5, that ultimately determine whether a height or a shape interpretation is more likely in our experiments. This shows that the analysis of gesture form has to include non-profiled articulators that impact the visual presentation of gesture form.

We should also emphasize that the visual presentation depends not only on the hand shape but also on the angle with which an observer looks onto the hand, an aspect not investigated in this study. For example, despite fingers 3-5 being slightly raised, they may still occlude the C-shape when the gesturing hand is viewed more from above. Follow-up experiments may investigate the impact of the viewing angle on gesture perception.

Finally, we should mention that the experiments presented here serve to demarcate ideal cases that are decontextualized. In a realistic speaking situation, the produced gestures usually interact with speech, which was outside of the scope of this investigation. For example, it is entirely possible that a fairly ambiguous gesture, somewhat in the middle of our two continua, is performed with the concomitant speech either emphasizing height or shape, and that this may lead people to attend to certain parts of the physical articulator. For example, it is surely possible to take a hand shape with fingers 3-5 raised and enforce a height reading by saying “the object was this high.” Alternatively, it is possible to verbally impose a shape reading on a height gesture despite fingers 3-5 occluding part of the C-shape. The interactions between speech and gesture are complex (e.g., Kendon, 2004, Ch. 8 & 9), but precisely because of this, there is value in looking at gesture alone to tease apart the relative contributions of gesture and speech. Future studies may investigate how the variables identified in the present experiments, INDEX CURVE and PINKIE CURL, interact with spoken language that implies either height or shape.

Summing up, we want to highlight some conclusions that we draw from our empirical data and analysis: First, it is important to distinguish between physical form and gesture form. Second, physical form reliably influences gesture form. In all tasks, curling in fingers 3-5 led to an increase in height gesture interpretations. Third, and finally, our results paint a complex picture of the conceptual differentiation between alternative gesture interpretations. Gestures are traditionally regarded as analog (McNeill, 1992), which certainly applies to some aspects of the gestures that were investigated in this paper: for example, a height gesture can indicate different heights on a continuous scale. On the other hand, the discrepancy between the gesture interpretations of a height or a shape gesture is difficult to describe in terms of a purely analog continuum: Even in the drawing task (where participants were not forced to respond in a binary fashion), participants tended to draw *either* an object aligned with the index finger and thumb, *or* an object filling out the C-shape. Unlike categorical speech perception, which shows thresholding effects, we did not find evidence for categorical boundaries with the variables considered in this paper (INDEX CURVE and PINKIE CURL). But a given individual appears to settle into drawing either one of two categories, i.e., either drawing an object aligned with the index/finger thumb axis, or not drawing an object along this axis. In conjunction with the relatively categorical contrast in production (§3), this suggests a clear role of categoricity in gestures that depict the spatial properties of objects.

Acknowledgments

We want to thank Gurjot Bains, Neekole Arcoda, Clarizza Arucan and Tania Lopez for help with coding the data. We want to thank Teenie Matlock for offering her lab to conduct parts of these studies. We want to thank Philipp Krecklow for modeling the virtual hands. Julius Hassemer specifically wants to thank his dissertation committee (Irene Mittelberg, Cornelia Müller, Klaus Willmes) for suggestions on an earlier stage of this project. This work was partially supported by NSF grant #1305196.

References

Andrén, M. (2010). *Children's Gestures from 18 to 30 months.* Dissertation, Lunds Universitet.

Beattie, G., & Shovelton, H. (2006). When size really matters: How a single semantic feature is represented in the speech and gesture modalities. *Gesture*, *6*(1), 63–84.

Bohannon, J. (2011). Social science for pennies. *Science*, 334, 307.

Bressem, J. (2013). A linguistic perspective on the notation of form features in gestures. In C. Müller, A. Cienki, E. Fricke, S. H. Ladewig, D. McNeill, & S. Teßendorf (Eds.), *Body-Language-Communication: An International Handbook on Multimodality in Human Interaction*. Berlin, Boston: Mouton de Gruyter.

Calbris, G. (1990). *The Semiotics of French Gestures*. Bloomington: Indiana University Press.

Calbris, G. (2003). From cutting an object to a clear cut analysis: Gesture as the representation of a preconceptual schema linking concrete actions to abstract notions. *Gesture, 3*(1), 19–46.

Calbris, G. (2011). *Elements of Meaning in Gesture*. Amsterdam/Philadelphia: John Benjamins.

Cienki, A. (2005). Image schemas and gesture. In B. Hampe (Ed.), *From perception to meaning: Image Schemas in Cognitive Linguistics* (pp. 421–442). Berlin: Mouton de Gruyter.

Emmorey, K., & Herzig, M. (2003). Categorical versus gradient properties of classifier constructions in ASL. In Emmorey, K. (Ed.), *Perspectives on classifier constructions in signed languages* (pp. 222-246). Mahwah: Lawrence Erlbaum Associates.

Gamer, M., Lemon, J., & Singh, I. F. P. (2012). irr: Various Coefficients of Interrater Reliability and Agreement. R package version 0.84.

Gentleman, R., & Lang, D. (2007). Statistical analyses and reproducible research. *Journal of Computational and Graphical Statistics, 16*, 1-23.

Hassemer, J. (2016). *Towards a Theory of Gesture Form Analysis. Imaginary forms as part of gesture conceptualisation, with empirical support from motion-capture data*. Dissertation, RWTH Aachen University.

Hassemer, J., Joue, G., Willmes, K., & Mittelberg, I. (2011). Dimensions and mechanisms of form constitution: Towards a formal description of gestures. In *Proceedings of Gesture and Speech in Interaction (GESPIN) 2011*. Bielefeld.

Holler, J., Shovelton, H., & Beattie, G. (2009). Do iconic hand gestures really contribute to the communication of semantic information in a face-to-face context? *Journal of Nonverbal Behavior*, *33*(2), 73–88.

James, W. (1890). *The principles of psychology. Vol. 2*. New York: Henry Holt.

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

Kita, S., van Gijn, I., & van der Hulst, H. (1998). Movement phases in signs and co-speech gestures, and their transcription by human coders. In I. Wachsmuth & M. Fröhlich (Eds.), *Gesture and Sign Language in Human-Computer Interaction. Proceedings of the International Gesture Workshop, September 17–19, 1997. Lecture Notes in Artificial Intelligence* (Vol. 1371, pp. 23–35). Berlin/Heidelberg: Springer.

Landis, J. R., & Koch, G. G. (1977). The measurement of observer agreement for categorical data. *Biometrics, 33*(1), 159–174.

Lempert, M. (2011). Barack Obama, being sharp: Indexical order in the pragmatics of precision-grip gesture. *Gesture*, *11*(3), 241–270.

Liberman, A.M., Harris, K.S., Hoffman, H.S., & Griffith, B.C. (1957). The discrimination of speech sounds within and across phoneme boundaries. *Journal of Experimental Psychology*, 54: 358–368.

Maechler, M. (2015). diptest: Hartigan's Dip test statistic for unimodality – corrected. R package version 0.75-7.

Mandel, M. A. (1977). Iconic devices in American Sign Language. In L. A. Friedman (Ed.), *On the Other Hand: New Perspectives on American Sign Language* (pp. 57–107). New York: Academic Press.

Mesirov, J. P. (2010). Computer science. Accessible reproducible research. *Science, 327*, 5964.

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

McNeill, D. (2005). *Gesture & thought*. Chicago: University of Chicago Press.

Mittelberg, I. (2010). Geometric and image-schematic patterns in gesture space. In V. Evans & P. Chilton (Eds.), *Language, Cognition, and Space: The State of the Art and New Directions* (pp. 351–385). London: Equinox.

Mittelberg, I., & Waugh, L. R. (2009). Metonymy first, metaphor second: A cognitive-semiotic approach to multimodal figures of thought in co-speech gesture. In C. Forceville & E. Urios-Aparisi (Eds.), *Multimodal Metaphor* (pp. 329–356). Berlin/New York: Mouton de Gruyter.

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

Müller, C. (2008). What gestures reveal about the nature of metaphor. In A. Cienki & C. Müller (Eds.), *Metaphor and Gesture* (pp. 219–245). Amsterdam/Philadelphia: John Benjamins.

Müller, C. (2014). Gestural modes of representation as techniques of depiction. In C. Müller, A. Cienki, E. Fricke, S. H. Ladewig, D. McNeill, & S. Teßendorf (Eds.), *Body – Language – Communication: An International Handbook on Multimodality in Human Interaction, Vol. 2* (pp. 1687–1702). Berlin/Boston: Mouton de Gruyter.

Parker, E. M., Diehl, R. L., & Kluender, K. R. (1986). Trading relations in speech and nonspeech. *Perception & Psychophysics*, 39(2), 129-142.

Peng, R. D. (2011). Reproducible research in computational science. *Science, 334*, 1226-1227.

Repp, B. H. (1982). Phonetic trading relations and context effects: New experimental evidence for a speech mode of perception. *Psychological Bulletin*, 92(1), 81-110.

R Core Team (2014). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL http://www.R-project.org/.

Sowa, T. (2006). *Understanding coverbal iconic gestures in shape descriptions*. Dissertation, University of Bielefeld.

Sprouse, J. (2011). A validation of Amazon Mechanical Turk for the collection of acceptability judgments in linguistic theory. *Behavior Research Methods*, 43, 155-167.

Stokoe, W. C. (2005/1960). Sign Language Structure: An Outline of the Visual Communication Systems of the American Deaf. *Journal of Deaf Studies and Deaf Education, 10*(1), 3–37.

Streeck, J. (2008). Depicting by gesture. *Gesture, 8*(3), 285–301.

Streeck, J. (2009). *Gesturecraft: The manu-facture of meaning.* Amsterdam: John Benjamins.

Talmy, L. (forthcoming.). *The targeting system of language*. MIT Press.

Winter, B., Perlman, M., & Matlock, T. (2013). Using space to talk and gesture about numbers: Evidence from the TV News Archive. *Gesture, 13*(3), 377–408.

Zwitserlood, I. (2012). Classifiers. In R. Pfau, M. Steinbach, & B. Woll (Eds.), *Sign language: An international handbook*. Berlin/Boston: De Gruyter.