# Gestures of the abstract

## Do speakers use space consistently and contrastively when gesturing about abstract concepts?

Fey Parrill and Kashmiri Stec
Case Western Reserve University, Cleveland, United States / Bang & Olufsen

Speakers perform manual gestures in the physical space nearest them, called *gesture space*. We used a controlled elicitation task to explore whether speakers use gesture space in a *consistent* way (assign spaces to ideas and use those spaces for those ideas) and whether they use space in a *contrastive* way (assign different spaces to different ideas when using contrastive speech) when talking about abstract referents. Participants answered two questions designed to elicit contrastive, abstract discourse. We investigated manual gesture behavior. Gesture hand, location on the horizontal axis, and referent in corresponding speech were coded. We also coded contrast in speech. Participants' overall tendency to use the same hand ($t(17) = 13.12$, $p = .001$, 95% CI [.31, .43], $d = 2.53$) and same location ($t(17) = 7.47$, $p = .001$, 95% CI [.27, .47], $d = 1.69$) when referring to an entity was higher than expected frequency. When comparing pairs of gestures produced with contrastive speech to pairs of gestures produced with non-contrastive speech, we found a greater tendency to produce gestures with different hands for contrastive speech: ($t(17) = 4.19$, $p = .001$, 95% CI [.27, .82], $d = 1.42$). We did not find associations between dominant side and positive concepts or between left, center, and right space and past, present, and future, respectively, as predicted by previous studies. Taken together, our findings suggest that speakers do produce spatially consistent and contrastive gestures for abstract as well as concrete referents. They may be using spatial resources to assist with abstract thinking, and/or to help interlocutors with reference tracking. Our findings also highlight the complexity of predicting gesture hand and location, which appears to be the outcome of many competing variables.

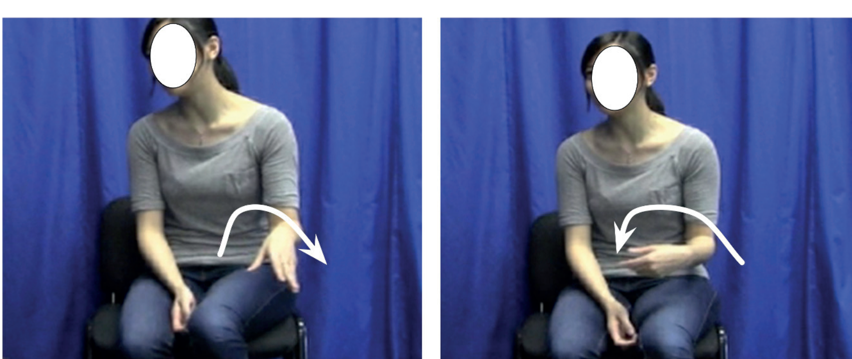
**Keywords:** gesture, referential space, spatial cognition, embodiment, multimodality

*Pragmatics & Cognition* 24:1 (2017), **33–61.** DOI 10.1075/pc.17006.par
ISSN 0929-0907 / E-ISSN 1569-9943 

## 1. Introduction

When using the hands to gesture, speakers use *gestural space*, the physical space surrounding a speaker's body, to locate their referents. When speakers gesture about concrete referents, the locations of the referents within gestural space are consistent. For example, after seeing a photo of a cat and a mouse, manual gestures co-occurring with speech about those objects will locate the cat in one place, the mouse in another, and those locations will be isomorphic to their locations in the photo. Research also suggests that spatial location for concrete referents can be used consistently for discourse purposes: A speaker might assign a space to a referent based on that referent's location in a photo, video, or remembered event, and that space will be used for the referent consistently (So, Kita & Goldin-Meadow 2009, Stec & Huiskes 2014, Perniss & Özyürek 2015). This consistent use of gestural space is thought to have both cognitive (Ping & Goldin-Meadow 2010) and communicative benefits (Stec & Huiskes 2014, Perniss & Özyürek 2015).

Do such patterns hold when the referent of a manual gesture is abstract? That is, when talking about an idea rather than a concrete object, do speakers assign the idea to an area of gestural space and use that space in a referentially consistent way? Certainly speakers do assign abstract referents to spatial locations – examples are commonplace. In the brief segment in Figure 1, a speaker produces a manual gesture that moves from a center rest position to the left of her body. She then produces a second manual gesture moving back to the center space. Without knowing what she was talking about, we might assume these gestures refer to concrete objects, perhaps a cat and a mouse. But rather than being a concrete object, the referent of the gestures is the abstract notion of *something good*, as described in English. The speaker's gestures locate good things in two positions: *Good things you do* to her left, and *good things happening to you* in the center, right in front of her.



Or if you do something good... something good is supposed to happen to you.

**Figure 1.** Example of Abstract Spatial Gesture

A few brief points about gesture for those readers who are less familiar with this area of research: McNeill (1992) defines gestures as "spontaneous movements" of the "arms and hands that are usually synchronized with the flow of speech" (p. 11). In this paper, we restrict ourselves to the prototypical use of the term, namely manual gestures or simply gestures. Gestures are produced both within gestural space (henceforth space) and within the context of a gesture phrase ("what we would intuitively call a gesture", McNeill 2005: 31). Gesture phrases are comprised of the following:

- An optional preparation phase, during which the hand moves from rest position into the gesture space where it can begin the stroke;
- An optional pre-stroke hold, in which the hand is temporarily held in place;
- An obligatory stroke phase is where the gesture occurs, conveying meaning while obviously produced with effort and focused energy;
- An optional post-stroke hold, in which the hand is temporarily held in place;
- An optional retraction phase, during which the hand either returns to rest or prepares for the next preparation phase (McNeill 1992).

By definition, the core component of a gesture is the stroke, and a core feature of the stroke is the physical location (or space) in which it occurs. Therefore, in this paper we investigate the stroke phase of each of the gestures of interest, as well as their spatial locations.

The phenomenon of gesture for abstract referents is well documented and is often described as *metaphoric* gesture (for European languages see, e.g., Sweetser 1992, Sweetser 1998, Cienki 1998, Cienki & Müller 2008, McNeill 1992, 2008, Müller 2004, 2008, Parrill 2007, Parrill & Sweetser 2004, Chui 2011 and for other languages see, e.g., Nuñez & Sweetser 2006, Cooperrider & Nuñez 2009, Le Guen & Pool Balam 2012). A number of researchers have also suggested that space is structured by a speaker during the production of abstract discourse. For example, McNeill (1992, 2000) argued that speakers re-index entities (whether concrete or abstract) by repeating gestures or elements of gesture, his notion of a *catchment*. McNeill and Pedelty (1995) point out that speakers set up semiotic spaces (p. 65), and distinguish between concrete space (where gesture location comes from the real world or the story world) and referential space (where space is used to refer to elements from the narrative, including abstract entities).

These studies have provided detailed analyses of individual examples (see Chui 2011 for an example of this approach), or have focused on consistent uses of space within a single semantic domain (e.g., spatial gestures for time, see Cooperrider & Núñez 2009 for an example). Drawing on naturalistic corpora or controlled elicitations, these researchers have established the existence of the phenomenon and, in cases such as the domain of time, have established some clear parameters of use. However, a systematic exploration across multiple domains is still needed, as a number of interesting questions remain open. Specifically, after setting up an association between a space and a referent, does a speaker continue to use that association? If the speaker in the example above talks about doing something good again, does she gesture to her left again? And how do these individual loci fit into a larger use of the space around her body? Does she use different locations contrastively in a larger discourse space? For example, to return to Figure 1, while *good happening to her* has been gestured at her location, and this contrasts with *good she does* to her left, if she talks about bad things, will she gesture to the right? In short, we know relatively little about the robustness, consistency, or elaboration of the use of space to represent abstract referents.

To explore this question, we present data from a controlled elicitation in which multiple speakers respond to two questions designed to elicit contrastive, abstract reference in gesture. We ask three research questions.

1. First, do participants use space in a *consistent* way? That is, do they assign a space to an idea and use that space when they refer to that idea? For example, a speaker might consistently use the left side of space to refer to good things, and also use the left side of space consistently to refer to *bad* things.
2. Second, do participants use space in a *contrastive* way? That is, do they assign different spaces to different ideas and use those spaces to contrast ideas? To clarify the difference, the previous example (good things located to the left; bad things also located to the left) demonstrates a use of space that is consistent but not contrastive. Using the left side of space for good and the right for bad, on the other hand, would indicate a contrastive use of space.
3. Third, do abstract referents have "predetermined" spatial locations? For example, if a speaker is talking about the abstract domains of research and teaching, is there any reason to think one would be on the left and the other on the right?

We know that gestures referring to concrete objects often derive their spatial locations from some actual location (in a photo, video, memory, real world scene…). However, we lack similar spatial knowledge for gestures referring to abstract concepts. It is not an entirely unexplored area, however. Research on abstract gestures, often called metaphoric gestures, suggests that abstract space may in fact come with some pre-existing structure. Metaphoric gestures reflect *conceptual metaphors*, or the understanding of an abstract domain in terms of another, concrete domain (Lakoff & Johnson 1980, 1999, Lakoff 1993). Conceptual metaphors have been shown to reflect underlying spatial schemas that manifest themselves in gesture (Sweetser 1992, 1998, Cienki 1998, Lakoff 2008, Müller 2008, Chui 2011). Cases well supported by research include gestures arising from the conceptual metaphor

TIME IS SPACE (Nuñez & Sweetser 2006, Cooperrider & Núñez 2009, Matlock et al. 2011, Le Guen & Pool Balam 2012). Gestures reflecting this conceptualization treat time as spatial expanses or locations, and often set up a lateral axis in which, for users of left-to-right writing systems, the past is gestured on the left, the present in the center, and the future on the right. Cultural differences of course apply, as the reverse appears to be true for users of right-to-left systems (Casasanto & Bottini 2014) and Mandarin speakers may use a vertical axis (Boroditsky 2001, Chui 2011). A second well-documented case is an association between *valence*, defined as positive or negative emotional appraisal, and handedness (Casasanto & Jasmin 2010). In this conceptualization, the dominant side (the left for left handers, the right for right handers) is associated with good and the non-dominant side with bad. Casasanto and Jasmin (2010) found that left-handers tended to gesture with the left hand for good, and right-handers tended to gesture with the right hand for good. This association appears to apply both to hand used, and to side of space in general (Casasanto & Henetz 2012). Other cases have been discussed as well, such as the structuring of the vertical dimension based on the conceptual metaphor GOOD IS UP (Lakoff & Johnson 1980). Gestures reflecting this conceptualization place a more positive concept higher in space than a more negative one (Cienki 1998). Thus, to restate our third research question, do the locations of the gestures we see in this corpus match the locations in this literature on metaphoric gestures?

Why do the answers to any of these questions matter for our understanding of the ways that humans think and communicate? Consistent use of space for concrete referents is thought to allow speakers to offload cognition onto the environment, freeing cognitive resources such as working memory and attention for communication or problem solving (McNeill & Pedelty 1995, So, Kita & Goldin-Meadow 2009). Further, consistent use of space may help listeners with their communicative task. While speakers may not be consciously aware of all co-speech gestures, they do attend to them (Gullberg & Kita 2009) and absorb information from them (Beattie & Shovelton 2002, 2005). A recent ERP study suggests that comprehenders do integrate abstract pointing gestures as they process language (Gunter, Weinbrenner & Holle 2015). Gunter and colleagues found a larger N400 (a measure of language processing that tends to be larger for unexpected or hard-to-process phenomena) for gestures that referred to an entity in a new location, as compared to gestures referring to that entity in an established location. They take this pattern as support for the claim that incongruent pointing gestures impair processing. Furthermore, as Stec and Huiskes (2014) have pointed out, the use of *referential space* (in the McNeill & Pedelty 1995 sense) is not driven by the speaker alone, but is negotiated by all participants in the discourse in a way that unfolds over time. It is a product not just of the speaker's cognition, but of the interaction.

In short, if we find that speakers structure space in predictable ways when talking about abstract ideas, and produce gestures that are spatially consistent and contrastive, speakers may be using spatial resources to assist with abstract thinking. If the relationship between gestures and their abstract referents is more transient, there is less reason to think speakers are capitalizing on space to organize their thoughts and communicate them.

Our final research question (Are abstract gestures spatially organized in ways that match predictions in the literature?) is important because the conceptual structuring of space around abstract concepts may represent both a competing pressure on a tendency to use space consistently and contrastively, and an explanation for the choices speakers make. In the case shown in Figure 1, the speaker used a left-right axis to represent good things one does (left) and good things happening to one (center). But previous work has suggested that speakers locate good things on their dominant sides, her right side in this case. Why does this speaker (who is right handed) place good things to the left and then center? The explanation may be that there is also *temporal* structuring to this example. The good things one does happen in the past relative to the good things that happen to one (*supposed to happen to you* = in the future, as a consequence of doing good). Metaphoric TIME IS SPACE gestures made by speakers of English locate the past to the left and the present in the center (ego's position), which is consistent with her gestural pattern, thus potentially explaining why she used these locations.

In short, the question of consistent and contrastive uses of abstract space necessarily involves exploring how abstract space might come pre-structured. These questions about consistency, contrast, and semantic structuring of space require answers. However, many obstacles arise when attempting to answer them. Before describing our study, we consider some of these challenges.

First, speech-gesture pairs are semiotically complex, and for the data described here, often create a *composite signal* in the sense of Clark (1996). That is, while saying *good things* and producing a gesture in a location, the speech is symbolic, while the gesture serves an indexical function (associates a location and an idea). The complexity of each pair and the different ways that gesture can index is beyond the scope of this study, but we refer readers to Clark (2003) and the papers in Kita (2003) for deeper discussions.

Second, there is no established definition for consistency in gesture and no established way of calculating it. In theory, consistency would be present if a participant produced the majority of gestures for a particular referent in one location or with the same hand or hands (e.g., always gesturing with both hands for *good things*). However, deciding whether a speaker is truly using gesture space consistently must refer to some chance distribution in order to be generalizable over speakers. It is unclear what "chance" frequency should be for a given gesture location. Outcomes (locations) aren't all equally likely. People generally tend to gesture in the center space (see, e.g., McNeill 1992, Priesters 2013). In addition, participants don't produce the same number of gestures for any referent, and don't use the available spaces equally often. Again, what is chance?

Third, we assume that because moving an articulator is effortful, speakers will continue to use the same hand and the same space unless motivated to change. But there are multiple competing principles in play at certain moments of the discourse. Imagine a speaker uses her right hand to gesture while talking about *good things*, then says something contrastive about those good things (e.g., good things in the past and good things in the future). While consistency might push her towards gesturing with her right hand (because the referent is still *good things*), contrast might push her towards using her left hand.

Further, contrast is only one of several reasons why speakers might be motivated to use different spaces. As discussed above, the semantic structuring of space is thought to be a powerful motivator for gesturing in different locations. There may be cases where different spaces are used not because of contrast, but because of an association between positive and dominant side, or the introduction of a hypothetical situation. There may also be cases where contrast appears in speech but hand or location does not change, because of larger discourse structuring principles.

In summary, the location in which a person produces a gesture is the outcome of many variables: Her overall tendency to gesture in that particular spatial location, the location of her previous gesture, the hand used for her previous gesture, the referent of the gesture, and the discourse level topic. The problem is considerably thornier for abstract reference than it is for concrete reference. We still believe it is worth addressing. We also suggest that an additional contribution of this study is a methodological beginning for these complex questions.

## 2. Method

### 2.1 Participants

Nineteen undergraduate students (13 women, mean age 20.11, range 18–26) from a private university in the northeastern United States took part in the study for either $10 or course extra credit. All were native speakers of English, seven were also fluent in another language. All reported that they were right handed.[1]

1. One reviewer notes that the Edinburgh Inventory is a better metric for lateral dominance than self-reporting. We agree that future studies investigating the organization of gesture space should include it as part of the experimental procedure but elected not do so as we were duplicating previous methods.

## 2.2 Procedure

Following informed consent, participants were seated at a computer where the instructions and prompts for the study were presented using Superlab (Cedrus Corporation). Participants received the following prompt: *We would like your opinion on the following question. Think about this question until you are ready to share your opinion, then press the space bar*. Participants then saw one of the following two questions, in random order: *Would you rather lose your hearing or your sight*? Or *Do you think the saying "what goes around comes around" is true?* Participants were then prompted, *Please go to the recording area and tell the experimenter what you think and why, giving a detailed answer*. After responding to the first question, participants returned to the computer and received the same prompt followed by whichever question they hadn't already received, and then responded to it. Participants were seated at a 90 degree angle to the experimenter, because the location of speaker and addressee is known to influence gesture for concrete spatial referents (Özyürek 2002). Participants also completed several additional tasks and measures not relevant for the current analysis.

The two questions given to participants were designed to generate contrastive discourse. Each requires the participant to consider two possibilities and express an opinion. In addition, both were expected to elicit abstract concepts of valence. In the case of the first question (hearing/sight), participants have to choose a preferred (positive) and dispreferred (negative) situation. In the case of the second question (*what goes around comes around*), participants respond by talking about good and bad actions and consequences, and also express a positive or negative opinion about the truth of the statement.

## 2.3 Coding

Speech and gesture coding was carried out in ELAN (Crasborn & Sloetjes 2008). ELAN tiers and controlled vocabulary are shown in the Appendix. Speech was transcribed. All manual gestures produced by participants were annotated and sorted into the following categories (after McNeill 1992, 2005): concrete iconic (features of the hand or body action map onto features of the stimulus event), metaphoric iconic (gestures that depict abstract content such as ideas or emotions, see Parrill & Sweetser 2004), deictic (gestures that point or locate objects in space), beat (purely rhythmic gestures). Further, concrete iconic gestures were coded as one of four types (also following McNeill 1992): iconic observer viewpoint path for path, iconic observer viewpoint path for shape, iconic observer viewpoint shape for shape, and iconic character viewpoint. Many gestures have more than one of these semiotic properties (e.g., an iconic gesture often deictically locates an object in space, and has a rhythmic component), so we coded using a decision tree. If a gesture had only a rhythmic component, it was treated as a beat. If it had rhythm and deixis, a deictic, etc. Each gesture was counted in only one category. We expected to see primarily metaphoric iconic gestures given the abstract nature of the discourse.[2]

Gestures were coded as left hand, right hand, or both. In order to be coded as a two handed gesture, both hands needed to execute a stroke phase (that is, gestures in which one hand is holding were not considered two handed). The spatial location of the stroke phase of all gestures was coded on the horizontal axis. Gestures were coded as left if the gesture occurred outside of the left hip or further away from the torso, right if the gesture occurred outside of the right hip or further away from the torso, occupying both left and right space (only possible for two handed gestures, in which case the criteria for left and right were applied in tandem), or center if the gesture occurred shoulder width or closer to the center of the torso.

Manual coding of the spatial location of gestures is challenging. As Priesters (2013) has noted, it is also challenging to do using technologically advanced tools. Priesters used motion capture to examine where in space participants gestured in a quasi-conversational setup. He also asked whether there were correlations between gesture location and gesture function (essentially asking about gesture consistency, though his analysis did not specifically explore metaphoric structuring of space). His work illustrates the problems inherent in studying this phenomenon. Due to the intensive nature of data analysis and coding, as well as the small size of his coded sample ($n = 4$), Priesters was not able to draw any strong conclusions. His main finding was that participants had idiosyncratic gesturing patterns. While we might have used a much more complex spatial coding system, we were dubious about the possibility of reliably capturing important aspects of location and trajectory. In short, this simple system (which we note ignores the vertical and sagittal axes entirely) is a first step. We discuss the coarse grain of spatial coding in the conclusions.

To code gesture referents, we developed a set of categories that emerged from the speech data, essentially those responses that were most frequent and relevant to the research questions. For the *hearing/sight* prompt, those categories were hearing, sight, related to the option the participant preferred (e.g., sight if the participant chose to hypothetically lose her sight), or other.[3] For the *goes around* prompt, those

2. We chose to use McNeill's coding scheme both because of its ubiquity in gesture research and so that our results would be directly comparable to those from Casasanto and Jasmin (2010). As one reviewer points out, this decision necessarily precludes categories for discourse or pragmatic gestures (e.g., Bavelas & Chovil, 2000) which frequently indicate the speaker's attitude or stance towards the discourse topic, classifying them instead as metaphoric iconic gestures.

3. There was also a category related to the option the participant did not prefer – e.g., hearing if the participant chose sight, but this did not occur in the dataset, so will not be discussed further.

categories were good things, bad things, a notion of cyclicity or a sequence of related events, agreement with the phrase, disagreement with the phrase, or other. Referents were placed into one of these categories based on co-occurring speech, and also the discourse level topic. For example, one participant generated a long list of cultural practices we use to make ourselves feel good, all of which were coded as *good* because of the larger connecting theme. *Other* was used only if the utterance could not be assigned to any of the previous categories (for example, *as I said before*). Thus, the coding was temporal rather than based on semantic features of the gestures (due in part to the fact that most gestures were palm up open hand gestures). It is also important to stress that while a gesture could in theory belong to multiple categories at once (e.g., a good thing that was part of a cycle of events), coders were asked to place each gesture into only one category. Future research could explore this multifunctional use of gesture.

### 2.4 Semantic coding

We were particularly interested in cases where an association between valence and dominant side was present, and cases where speech suggested the conceptual metaphor TIME IS SPACE was being used. To explore the association of good and dominant side or hand, we adapted the method of Casasanto and Jasmin (2010). The ELAN speech output was divided into clauses. Following Casasanto and Jasmin, the valence of each clause was coded as "…either positive, negative, neutral, or indeterminate (i.e., ambiguous or mixed valence)" (p. 4). We selected the clauses coded as positive or negative ($N = 337$), then determined which of those clauses was accompanied by a gesture or gestures, using the full ELAN output ($N = 199$). We eliminated any clauses accompanied by two-handed gestures ($N = 91$). The good/dominant analysis, then, used just those clause-gesture pairs where speech was about something positive or negative, and a one-handed gesture was performed ($N = 108$).

For the TIME IS SPACE analysis, we coded all clauses as referring to the past (e.g., *I did something good first*), the present (*right now*), the future (e.g., *and, like, it can benefit you later in life*), or mixed (e.g., *it's kind of, like, a greater cycle*). Clauses accompanied by a one-handed gesture or gestures were used for the TIME IS SPACE analysis ($N = 206$).

## 2.5 Contrast coding

Contrast in speech is defined for this study as a situation in which two propositions are opposed, or in which a presupposed proposition is negated (Ford 2000). Contrast is expressed using both syntactic and lexical means, including clefting, fronting, and coordination (Biber et al. 1999). We elected to focus on lexical items that are used for contrast, and searched the transcript for the expressions of contrast and concession listed in Table 1, an approach informed by the work of Hinnell and Rice (2016). This set of expressions is not complete, nor does it contain two coordinating conjunctions often used for contrast, *and* and *or*. We excluded those conjunctions because they are highly polysemous, and our study is intended to be a preliminary step in approaching the question of contrast in gesture.

**Table 1.** Linguistic expressions of contrast

| Expression | Count | Expression | Count |
|---|---|---|---|
| Alternatively | 0 | In contrast | 0 |
| Although / even though | 4 | In spite of | 0 |
| Anyhow | 0 | Instead | 0 |
| Anyway | 0 | Neither … nor… | 0 |
| At any rate | 0 | Nevertheless | 0 |
| Besides | 0 | On the contrary | 0 |
| But | 53 | On the one hand… | 0 |
| By comparison | 0 | On the one hand…on the other hand | 0 |
| By contrast | 0 | On the other hand… | 0 |
| Conversely | 0 | Rather (than) | 20 |
| Despite | 0 | That being said | 1 |
| Even though | 2 | Whereas | 4 |
| However | 2 | Yet | 0 |
| In any case | 1 | | |

## 2.6 Reliability

To establish reliability of coding for gesture, one coder coded all data, and a second coder coded 4 randomly selected participants (21% of the final dataset). Reliability is expressed as percent agreement and Cohen's kappa, as follows: Gesture type: 93%, .82; gesture horizontal axis: 94%, .90; gesture hand: 99%, .99; gesture referent: 90%, .88; clause valence: 93%, .88; clause temporal category: 94%, .98; contrast in speech: 100%.

## 3. Results

To answer our research questions, we performed a quantitative analysis of our data. Our results will be presented as follows. We first present a general description of the data. We next describe analyses of consistency. For reasons described below, we carry out three separate consistency analyses. Next we carry out an analysis of contrast. Finally we describe analyses for the metaphoric organization of space. We conduct separate analyses for the two domains of interest.

### 3.1 General description of the data

The mean gesture rate (number of gestures divided by number of clauses for each participant, averaged over participants) was 10.18 (SD 7.78). Participants produced an average of 38.38 clauses (SD 22.65). Table 2 shows the mean proportion of each gesture type. Mean proportions are calculated as the number of each type of gesture divided by the total number of gestures, averaged over participants, and are used to account for large differences in individual gesture rates. Table 2 also shows mean proportions for hand used, gesture location, and gesture referent.

**Table 2.** Gesture type and gesture distributions, mean proportions

| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| **Type** | **Beat** | **Deict** | **Emblem** | **Icon** | **Metaphoric** | | | | |
| | .15 (.15) | .03 (.04) | .03 (.06) | .04 (.06) | .75 (.12) | | | | |
| **Hand** | **BH** | **Left** | **Right** | | | | | | |
| | .48 (.28) | .22 (.25) | .29 (.32) | | | | | | |
| **Loc** | **Center** | **Left** | **L/R** | **Right** | | | | | |
| | .60 (.31) | .16 (.15) | .10 (.12) | .14 (.24) | | | | | |
| **Ref** | **Agree** | **Dis** | **Cycle** | **Bad** | **Good** | **Hear** | **Sight** | **Pref** | **Other** |
| | .03 (.04) | .06 (.10) | .18 (.15) | .15 (.15) | .15 (.19) | .18 (.13) | .17 (.21) | .05 (.08) | .04 (.11) |

### 3.2 Consistency analysis

Before conducting consistency and contrast analyses, we eliminated one participant who performed only one gesture: Consistency and contrast are not possible with a single gesture. For the consistency analysis (but not the contrast analysis), we also eliminated gestures with the referent "other" from the analysis (8 gestures total). This is because instances of "other" are not homogenous, and thus can't be treated as the same referent (e.g., two instances of other, *as I said before* and *so an example*, are not the same referent).

As discussed above, there is no existing definition of consistency in gesture location, nor is there an established procedure for calculating it. We therefore carried out three separate analyses that highlight different dimensions of consistency. Consistency will be defined as two gestures with the same referent produced in the same space and/or with the same hand. We chose to consider not just location but also hand used because, as discussed in our introduction, there are many reasons why a gesture might "end up" in a particular location. Considering hand used may prove to be a fairer measure of consistency. As a first pass, we simply asked what proportion of participants produced at least one pair of consistent gestures. All but one participant did so.

### 3.2.1 *Overall consistency*

Our next step was to attempt to obtain an overall picture of consistency by assessing the extent to which a speaker gestured with a particular hand or in a particular location for a particular referent. We will describe this analysis beginning with hand used, then location. An example of this process is given in the Appendix to clarify the calculations. All gestures produced by each participant were used for this analysis.

We first created a contingency table for each participant (a table showing frequencies for gesture referents according to hand used). For each gesture referent, we identified the most frequently used hand, and asked what proportion of gestures with that referent were made with that hand. This step captures the idea that consistency is present when a speaker uses the same hand a majority of the time when talking about a particular referent. We did this calculation for all referents for a particular participant, then averaged across those results within a speaker to yield an average measure of consistency for each participant. Next, we created a contingency table of expected frequencies for each person. Expected frequencies are calculated by multiplying row totals (hands used) by column totals (gesture referents), then dividing by total number of gestures for that participant. These expected values can be averaged over participants in the same fashion, yielding an expected frequency for each participant. This process allowed us to compare a participant's actual tendency to use a hand for a referent the majority of the time to the expected frequency. If participants do in fact tend to use the same hand for the same referent, their actual frequency should be higher than expected frequency. Again, an example of this process can be found in the Appendix.

We used a paired t-test (see the Appendix for our explanation of this choice) to compare these mean proportions. We found that the mean proportion for actual frequency was significantly higher than the mean proportion for expected frequency for using the same hand for the same referent ($t(17) = 13.12$, $p = .001$, 95% CI [.31, .43], $d = 2.53$). We performed the same analysis for the relationship between gesture referent and gesture location. For location, the mean proportion for actual frequency was significantly higher than the mean proportion for expected frequency ($t(17) = 7.47$, $p = .001$, 95% CI [.27, .47], $d = 1.69$). This analysis suggests that a participant's overall tendency to use a particular hand or location when referring to an entity is higher than would be expected if gestures for a referent were distributed across hand and spaces used by that participant in a balanced way.

**Table 3.** Mean actual and expected frequencies for hand and location

| | **Hand** | **Location** |
|---|---|---|
| Actual frequency | .83 (.14) | .81 (.15) |
| Expected frequency | .45 (.16) | .44 (.27) |

### 3.2.2 *Consistency in temporally sequential gestures*

The strength of the previous approach is that it provides an overall measure of participants' tendency to use the same hand and/or location for a referent. However, it has at least two weaknesses. First, it assumes the most frequent hand used for a given referent represents that speaker's "choice" for that referent. If a speaker uses the right hand five times for good things, and the left hand four times for good things, it is not really the case that the right hand has been assigned to good things – both hands are being used consistently. The proportional calculation addresses this problem, but does not entirely solve it. Second, this approach ignores the variable of time. Gestures with the same referent may be very far apart in time, and while such gestures would absolutely be considered a consistent use of space, gestures that are far apart in time might be more likely to be produced in different locations or with different hands. (Because, as discussed, there are many reasons why speakers change location and hand.) For both these reasons, the analysis just described – while being a legitimate measure of consistency – may underrepresent consistency.

Our final analysis attempted to build time back into the picture, by focusing on pairs of temporally sequential gestures. While there are many pressures causing a speaker to use a different hand or location, if gesture hand and gesture location are used to index reference, sequential gestures with the same referent should be more likely to be produced with the same hand and/or occur in the same location than sequential gestures with different referents.

For this analysis, each participant's data was sorted into two sets of pairs: Pairs of temporally sequential gestures with the same referent ($N = 107$), and pairs of temporally sequential gestures with different referents ($N = 76$). We simply started at the beginning of the person's response and paired gestures that occurred within 8 seconds of each other (the average window for these data). Twenty-one gestures were eliminated from the dataset because they were not in a sequential pair. We then asked how often different hands and/or different spaces were used for those two categories of gestures. Table 4 shows mean proportions for pairs of gestures produced with different hands, or occurring in different spaces, as a function of whether the referent was the same or different. Note that if an average of 25% of participants used different hands when the referent was the same, this entails that an average of 75% of participants used the same hand. A paired t-test showed no statistically significant difference between the mean proportion of gestures made with different hands for same versus different referents: $t(17) = .99$, $p = .34$. A paired t-test showed no statistically significant difference between the mean proportion of gestures made in different spaces for same versus different referents: $t(17) = .24$, $p = .81$. It thus appears that while participants did generally use the same hand when producing pairs of gestures with the same referent (i.e. they only used different hands for an average of 25% of pairs), they also used the same hand when producing pairs of gestures with *different* referents (i.e. they only used different hands for an average of 33% of pairs).

**Table 4.** Paired gesture analysis: Hand and location for same and different referent pairs

| | **Same referent** | **Different referent** |
|---|---|---|
| Different hands | .25 (.30) | .33 (.35) |
| Different location | .43 (.34) | .46 (.43) |

### 3.3 Contrast analysis

Our second research question is whether participants used space *contrastively* by assigning different spaces to different ideas. We first simply asked what proportion of participants produced at least one instance of contrastive gesture use, defined as two gestures with different referents, produced in different spaces or using different hands. All but one participant did this. We next identified all cases where contrast was present in speech (that is, the occurrence of the contrastive linguistic expressions listed in Table 1). If contrast occurred in speech, we asked whether gestures were produced for the two immediately contrasting elements (the propositions being opposed by the speaker). If so, we asked whether these pairs of gestures were made with different hands and/or in different spaces. Table 5 shows the mean number of contrastive expressions, the mean proportion of contrastive speech accompanied by gesture, and mean proportions for gestures produced with different hands, occurring in different spaces. Almost 60% of gesture pairs occurring with contrastive speech were produced with different hands, though there was quite a bit of variability across participants, as indicated by the large SDs. To determine whether different hands and spaces were more likely to be used when speech was contrastive, we randomly selected the same number of pairs of gestures accompanying non-contrastive speech for each participant and asked whether different hands and different spaces were used for those gestures. These mean proportions are presented in Table 5 as well. A paired t-test showed a statistically significant difference between the mean proportion of gestures made with different hands for contrastive versus non-contrastive speech: $t(17) = 4.19$, $p = .001$, 95% CI [.27, .82], $d = 1.42$. A paired t-test showed no statistically significant difference between the mean proportion of gestures made in different locations for contrastive versus non-contrastive speech: $t(17) = .79$, $p = .44$. These patterns indicate that while contrastive speech was more likely to be accompanied by a change in gesture hand, it was not more likely to be accompanied by a change in gesture location.

**Table 5.** Contrastive and non-contrastive speech comparison

| | **Contrastive speech** | **Non-contrastive speech** |
|---|---|---|
| Expressions of contrast | 4.83 (2.83) | – |
| With gesture | .73 ( .22) | – |
| Different hands | .58 ( .53) | .04 (.09) |
| Different spaces | .41 ( .40) | .27 (.53) |

*Note*. Expressions of contrast is a mean number, all other values are mean proportions, SDs in parentheses.

## 3.4 Spatial organization analysis

Our final question is whether abstract gestures were spatially organized in ways that match predictions in the literature. Specifically, did we see an association between positivity and the dominant hand or side of space (right, since all participants were right handed), or evidence of a horizontal timeline (namely past on the left, present in the center, future on the right)? We conducted separate analyses for the metaphoric domains of interest.

### 3.4.1 *Good and dominant side*

To explore the good on the dominant, bad on the non-dominant side pattern, we used the clause-based coding of valence described above. Recall that all clauses coded as negative or positive and accompanied by a one-handed gesture were used for the analysis (following Casasanto & Jasmin 2010). Because all our participants were right handed, we looked for a relationship between gestures made with the right hand and positive valence, and between gestures made with the left hand and negative valence. Table 6 shows the frequencies of gestures performed with the left and right hands, according to the valence of the co-occurring clause. A Chi-square test of independence suggests this distribution is not more likely than expected frequency ($\chi^2$ (1) = .01, $p$ = .85).

**Table 6.** Distribution of gesture hand according to valence of clause

| | Left | Right | Total |
|---|---|---|---|
| Negative | 22 | 30 | 52 |
| Positive | 22 | 34 | 56 |
| **Total** | **44** | **64** | **108** |

We also considered the relationship between valence and horizontal space. A gesture might be made with the left hand but to the right of the body, thus presenting a different, but motivated, picture of valence and dominant side. Table 7 shows these frequencies. Again, a Chi-square test of independence suggests this distribution is not more likely than expected frequency ($\chi^2$ (3) = .86, $p$ = .84).

**Table 7.** Distribution of gesture location according to valence of clause

| | Center | Left side | Right side | Total |
|---|---|---|---|---|
| Negative | 22 | 13 | 17 | 52 |
| Positive | 19 | 17 | 20 | 56 |
| **Total** | **40** | **44** | **64** | **108** |

### 3.4.2 *TIME IS SPACE analysis*

To explore the horizontal timeline (past is left, present is center, future is right), we used the clause-based coding of time described above. All clauses coded as past, present, future, or mixed and accompanied by a one-handed gesture were used for the analysis. (While most of these came from the *goes around* prompt, we included all clauses.) Table 8 shows the frequencies of gestures performed with the left and right hands, according to the temporal category of the clause. If participants are using the left hand for past and the right for future, we would expect to see this in the distribution. A Chi-square test of independence suggests this distribution is not more likely than expected frequency ($\chi^2$ (3) = 2.42, $p$ = .49). We conducted the same analysis for gesture location, since a speaker needn't encode TIME IS SPACE by changing hands, but can simply change locations. Table 9 shows the frequencies of gestures performed in center, left, and right space according to the temporal category of the clause. If speakers are gesturing in the left space for past, center for present, and right for future, we should see this pattern in the distribution. A Chi-square test of independence suggests this distribution is not more likely than expected frequency ($\chi^2$ (9) = 10.54, $p$ = .49).

**Table 8.** Distribution of gesture hand according to temporal category of clause

| | Left | Right | Total |
|---|---|---|---|
| Past | 6 | 6 | 12 |
| Present | 55 | 84 | 139 |
| Future | 19 | 29 | 48 |
| Mixed | 1 | 6 | 7 |
| **Total** | **81** | **125** | **206** |

**Table 9.** Distribution of gesture location according to temporal category of clause

| | Center | Left | Right | Total |
|---|---|---|---|---|
| Past | 5 | 4 | 3 | 12 |
| Present | 54 | 38 | 47 | 139 |
| Future | 26 | 10 | 12 | 48 |
| Mixed | 3 | 0 | 4 | 7 |
| **Total** | **88** | **52** | **66** | **206** |

## 4. Discussion

This study explored four questions about the relationship between gesture and abstract reference. Do speakers assign a hand or space to an idea and use that hand or space consistently when they refer to that idea? Second, do speakers use space in a contrastive way by assigning different hands or spaces to different ideas? Third, do abstract referents have "predetermined" spatial locations? Finally, are gestures spatially organized in ways that match predictions in the literature?

Our first analysis captured the idea that consistency is present when a speaker uses the same hand or location a majority of the time when talking about a particular referent. Speakers did use hand and location consistently according to this metric. They were more likely to use the same hand and the same location for a referent than would be expected if gestures were distributed in a balanced way. However, an analysis exploring the extent to which pairs of temporally sequential gestures for the same referent were produced with the same hand or in the same location found that our participants were equally likely to use the same hand and location for gestures with the same referent and for gestures with different referents. These findings are not contradictory. They reflect the fact that different approaches to consistency highlight different aspects of the phenomenon. An example from one speaker may help to clarify this point. Four of this speaker's temporally sequential gestures were as follows:

1. Right hand, gesture referent = sight
2. Both hands, gesture referent = hearing
3. Both hands, gesture referent = hearing
4. Right hand, gesture referent = sight

Our first approach to consistency captures the fact that the speaker is using right hand for hearing and both hands for sight for these four gestures. However, the second analysis paired gestures 1 and 2, 3 and 4. These pairs have different referents. Using temporally sequential gestures without regard to any larger structuring of the discourse corrects some of the weaknesses of the first approach, but means losing some cases of consistency. These two findings also reflect the fact that many variables interact in determining gesture hand and location. The second pattern – namely, a tendency to use the same hand or location for gestures with the same referent as well as for gestures with different referents – lends support to the idea that changing articulators is effortful. In summary, consistency is particularly hard to capture because it can exist both at a very local level in a pair of gestures and at a discourse level.

A contrastive use of gesture emerged when we considered pairs of gestures occurring with contrastive speech. Those pairs were more likely to be made with different hands than pairs of gestures occurring with non-contrastive speech. Contrast may therefore be a powerful force shaping gesture hand selection.

In this dataset, participants did not show any tendency to gesture in ways that indicated a metaphorical structuring of space when we considered overall patterns. First, our participants did not show an overall tendency to gesture with the dominant hand, or on the dominant side, for clauses with positive valence. We suggest two reasons for this difference between our findings and those of Casasanto and Jasmin (2010). There were far fewer gestures in our dataset (108) compared to the sample used for their study (763), and our sample comes from a larger number of individuals (19 vs. their 4). More important is the fact that our dataset had almost an equal balance between right- and left-handed gestures. In Casasanto and Jasmin's data, only 16% of the gestures produced by right-handed speakers were made with the left hand, whereas our participants (all right handers) produced 41% of gestures for negatively or positively valenced clauses with the left hand. There are many possible explanations for this difference in distribution, including discourse genre (our quasi-conversation vs. their political debate), speaker characteristics (our college students vs. their trained public speakers), or the position of the interlocutor (to the speaker's right for our data, versus a complex situation with debate opponent, moderator, real audience, camera, and television audience, for their data).

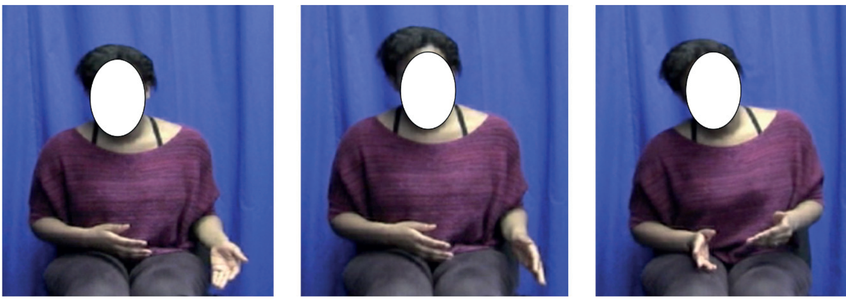
Similarly, we did not find that gestures for speech about the past tended to occur on the left side or be made with the left hand more often, nor did gestures for speech about the future tend to occur on the right side or be made with the right hand more often. Converging evidence does support both of these associations (good and dominant, left/right and past/future), so we do not see our results as being particularly strong counterevidence. Indeed, we saw many cases where we suspect the motivation for gesture hand or gesture location was driven by some of these conceptual structures. A case is discussed below (see Figure 2). Rather, we believe that these associations are only one conceptual pressure in the process of gesture hand or location selection.

Some of our findings are likely shaped by the level of detail at which we analyzed gesture. For instance, we did not do a deep analysis of the different semantic, syntactic or lexical items that were coded as the same referent in speech, which might be illuminating. Nor did we annotate gesture shape or orientation, and these variables are known to encode important semantic information. As an example, in the case shown in Figure 2, both gestures are made with the right hand in what would be coded as center space. However, the speaker changes palm orientation for the two different gesture referents. Finer-grained analyses might reveal more about speaker-specific patterns as well as general tendencies.

Does something to you... ...you get revenge on them

**Figure 2.** Change of palm orientation for change of reference.

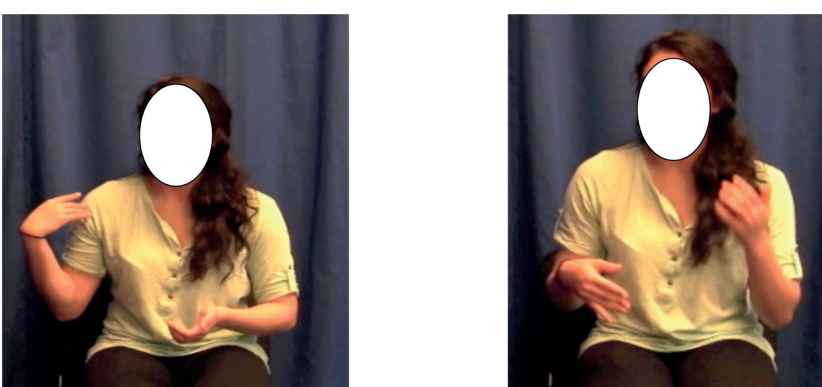
In an effort to model our analyses on existing approaches, the good/dominant analysis did not include two-handed gestures. These gestures can encode valence in rich and interesting ways. For example, Figure 3 shows a speaker who starts by saying *good things* while producing a two-handed gesture with the left palm up. She then interrupts herself and produces the same gesture with the speech *or bad* *things to good people*, and then produces a final two-handed gesture with the right palm up while saying and *good things to bad people*. While we can only speculate about why she interrupts herself, the repair may serve to keep the valence of good and bad things consistent with her dominant and non-dominant sides.



Good things* or BAD things to good people.... and good things to bad people

**Figure 3.** Two handed gesture for positive / negative valence.

Other examples occurred in which a change in the palm orientation of pairs of two-handed gestures appeared to serve some function, as in the cases in Figures 4 and 5. In both cases the speaker changes which palm faces down during the stroke phase of a two handed gesture. These balance gestures may be consistent with the dominant/good association, depending on which aspect of the gesture one focuses on.



Kind of, like, associate, like, a positive... or, like, negative thing

**Figure 4.** Two handed gesture for positive / negative valence.

Try to do good... even though it might not work out

**Figure 5.** Two handed gesture with change in palm orientation

Examples such as these illustrate the tension between the case study method, in which small numbers of examples are explored in order to understand a phenomenon, and the generalization-over-cases method used here, in which rich detail is ignored in order to make the coding tractable. Certain phenomena slip through the cracks because of the granularity with which we coded gesture, but that level of granularity was necessary in order to be able to address specific questions.

In addition to hand shape and orientation, we considered only the horizontal dimension, and did not code the vertical or sagittal axis. While we could have included these dimensions, it is unclear how we would have folded them into the analyses. To require that a gesture have the same location in three dimensions in order to be considered consistent would seem to be an overly strict criterion for consistency. Similarly, to consider a gesture that differed on any of the three dimensions as a different location might have been an overly lax criterion for contrast. We suggest that our level of analysis is a good fit for our questions, and hope that others will be inspired to correct the weaknesses of this approach.

## 5. Conclusions

We used a controlled elicitation task to explore whether speakers use gestural space in a consistent way (assign ideas to spaces and use those spaces for those ideas) and whether they use space in a contrastive way (assign different ideas to different spaces when using contrastive speech) when talking about abstract referents. We found support for both behaviors. An analysis of consistency that looked at overall tendency to associate a referent with a hand or space found that speakers do tend to do so. Speakers were more likely to produce gestures with different hands for contrastive speech compared to non-contrastive speech. However, gesture-speech association also revealed the complexity of predicting where and with which hand a speaker will gesture.

We believe that gesture hand and location are the outcome of a multidimensional space. Modeling this behavior will require more sophisticated tools than those we were able to employ. Indeed, gesture production might be fruitfully considered as a dynamical system (van Gelder 1998, Beer 2000). If gesture location is a trajectory through a state space, some of the dimensions modeled might include speaker's dominant hand, interlocutor location, current gesture referent, previous gesture referents, degree of contrast in speech, and many others. Such an approach might work well with Hostetter and Alibali's (2008) notion of a gesture threshold (the factors that determine whether or not a person gestures). If changing articulators is in fact, effortful, we might predict that the location and hand of a gesture will remain the same as the previous one unless the activity of some dimension (e.g., contrast in speech) exceeds a threshold. Such an approach also needs to mesh the physical and communicative dimensions of this system (see Kok et al. 2016). Our overall point is simply that this behavior is complex, and a complex analytic approach will likely be necessary in order to fully understand it.

What do our findings contribute to the study of how language is used? Most broadly, they illuminate the ways in which discourse is constructed using multimodal resources. We believe that the location and hand used in producing a gesture for an abstract referent is non-random, i.e., that it reflects something important about the speakers' cognitive processes. We suggest that consistent and contrastive use of gesture use allows speakers to save memory and attention resources by offloading cognition onto the environment, as has been suggested for concrete reference. A next step might be to ask whether contrastive gesture is impacted when speakers are only allowed to use one hand. Do they change location more frequently?

We also believe that listeners track gesture hand and location – though likely below the level of conscious awareness – and that a speaker's choice impacts comprehension. This claim can be tested using paradigms like those in Gunter, Weinbrenner and Holle (2015), but another approach might be to assess the extent to which an addressee builds a mental map of referent-location pairings for abstract referents, potentially as compared to concrete referents. Future research should investigate this possibility.

As Clark (2003) puts it, "[p]eople communicate […] by creating signs by which they mean things for others" (p. 245). This meaning-making process is multimodal, and complex. Our study provides an entry point to systematizing some of that complexity.

## References


Beattie, Geoffrey & Heather Shovelton. 2002. An experimental investigation of some properties of individual iconic gestures that mediate their communicative power. *British Journal of Psychology* 93. 179–192. doi:10.1348/000712602162526

Beattie, Geoffrey & Heather Shovelton. 2005. Why the spontaneous images created by the hands during talk can help make TV advertisements more effective. *British Journal of Psychology* 96(1). 21–37. doi:10.1348/000712605X103500

Bavelas, Janet Beavin & Nichole Chovil. 2000. Visible acts of meaning: An integrated message model of language in face-to-face dialogue. *Journal of Language and Social Psychology* 19(25). 163–194. doi:10.1177/0261927X00019002001

Beer, Randall. 2000. Dynamical approaches to cognitive science. *Trends in Cognitive Sciences* 4(3). 91–99. doi:10.1016/S1364-6613(99)01440-0

Biber, Douglas, Stig Johansson, Geoffrey Leech, Susan Conrad & Edward Finegan. 1999. *Longman grammar of spoken and written English*. Harlow: Pearson Education Limited.

Boroditsky, Lera. 2001. Does language shape thought? Mandarin and English speakers' conceptions of time. *Cognitive Psychology* 43. 1–22. doi:10.1006/cogp.2001.0748

Casasanto, Daniel & Roberto Bottini. 2014. Mirror reading can reverse the flow of time. *Journal of Experimental Psychology: General* 143(2). 473–479. doi:10.1037/a0033297

Casasanto, Daniel & Tania Henetz. 2012. Handedness shapes children's abstract concepts. *Cognitive Science* 36. 359–372. doi:10.1111/j.1551-6709.2011.01199.x

Casasanto, Daniel & Kyle Jasmin. 2010. Good and bad in the hands of politicians: Spontaneous gestures during positive and negative speech. *PLoS ONE* 5(7). e11805. doi:10.1371/journal.pone.0011805

Chui, Kawai. 2011. Conceptual metaphors in gesture. *Cognitive Linguistics* 22(3). 437–458. doi:10.1515/COGL.2011.017

Cienki, Alan. 1998. Metaphoric gestures and some of their relations to verbal metaphoric expressions. In Jean-Pierre Koenig (ed.), *Discourse and cognition: Bridging the gap*, 189–204. Stanford: CSLI Publications.

Cienki, Alan & Cornelia Müller (eds.). 2008. *Metaphor and gesture*. Amsterdam: John Benjamins. doi:10.1075/gs.3

Clark, Herbert H. 1996. *Using language*. Cambridge: Cambridge University Press. doi:10.1017/CBO9780511620539

Clark, Herbert H. 2003. Pointing and placing. In Sotaro Kita (ed.), *Pointing: Where language, culture, and cognition meet*, 243–268. New Jersey: Lawrence Erlbaum.

Cooperrider, Kensy & Rafael Núñez. 2009. Across time, across the body. *Gesture* 9(2). 181–206. doi:10.1075/gest.9.2.02coo

Crasborn, Onno & Han Sloetjes. 2008. Enhanced ELAN functionality for sign language corpora. In Onno Crasborn, Thomas Hanke, Eleni Efthimiou, Inge Zwitserlood & Ernst Thoutenhoofd (eds.), *Construction and exploitation of sign language corpora: 3rd workshop on the representation and processing of sign languages*, 39–43. Paris: ELDA.

Ford, Cecilia E. 2000. The treatment of contrasts in interaction. In Elizabeth Cooper-Kuhlen & Bernd Kortmann (eds.), *Cause, condition, concession, contrast: Cognitive and discourse perspectives*, 283–312. Berlin: Mouton de Gruyter. doi:10.1515/9783110219043.3.283

Gullberg, Marianne & Sotaro Kita. 2009. Attention to speech-accompanying gestures: Eye movements and information uptake. *Journal of Nonverbal Behavior* 33. 251–277. doi:10.1007/s10919-009-0073-2

Gunter, Thomas C., J. E. Douglas Weinbrenner & Henning Holle. 2015. Inconsistent use of gesture space during abstract pointing impairs language comprehension. *Frontiers in Psychology* 6(80). 1–10. doi:10.3389/fpsyg.2015.00080

Hinnell, Jennifer & Sally Rice. 2016. *"On the one hand…": Opposition and optionality in the embodied marking of stance in North American English*. Paper presented at the International Society for Gesture Studies, Paris.

Hostetter, Autumn B. & Martha W. Alibali. 2008. Visible embodiment: Gesture as simulated action. *Psychonomic Bulletin & Review* 15(3). 495–514. doi:10.3758/PBR.15.3.495

Kita, Sotaro (ed.). 2003. *Pointing: Where language, culture, and cognition meet*. Mahwah, NJ: Lawrence Erlbaum Associates.

Kok, Kasper, Kirsten Bergmann, Alan Cienki & Stefan Kopp. 2016. Mapping out the multifunctionality of speakers' gestures. *Gesture* 15(1). 37–59. doi:10.1075/gest.15.1.02kok

Lakoff, George. 1993. The contemporary theory of metaphor. In Andrew Ortony (ed.), *Metaphor and thought*, 202–251. Cambridge: Cambridge University Press. doi:10.1017/CBO9781139173865.013

Lakoff, George. 2008. The neuroscience of metaphoric gestures: Why they exist. In Alan Cienki & Cornelia Müller (eds.), *Metaphor and gesture*, 283–289. Amsterdam: John Benjamins.

Lakoff, George & Mark Johnson. 1980. *Metaphors we live by*. Chicago: University of Chicago Press.

Lakoff, George & Mark Johnson. 1999. *Philosophy in the flesh: The embodied mind and its challenge to Western thought*. New York: Basic Books.

Le Guen, Oliver & Lorena I. Pool Balam. 2012. No metaphorical timeline in gesture and cognition among Yucatec Mayas. *Frontiers in Psychology* 3. 1–15. doi:10.3389/fpsyg.2012.00271

Matlock, Teenie, Kevin J. Holmes, Mahesh Srinivasan & Michael Ramscar. 2011. Even abstract motion influences the understanding of time. *Metaphor and Symbol* 26. 260–271. doi:10.1080/10926488.2011.609065

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

McNeill, David (ed.). 2000. *Language and gesture*. Chicago: University of Chicago Press. doi:10.1017/CBO9780511620850

McNeill, David. 2005. *Gesture and thought*. Chicago: University of Chicago Press. doi:10.7208/chicago/9780226514642.001.0001

McNeill, David. 2008. Unexpected metaphors. In Alan Cienki & Cornelia Müller (eds.), *Metaphor and gesture*, 155–170. Amsterdam: John Benjamins.

McNeill, David, & Laura Pedelty. 1995. Right brain and gesture. In Karen Emmorey & Judy Reilly (eds.), *Language, gesture and space*, 63–85. Hillsdale, NJ: Lawrence Erlbaum.

Müller, Cornelia. 2004. *Metaphors: Dead and alive, sleeping and waking*. Chicago, IL: University of Chicago Press.

Müller, Cornelia. 2008. What gestures reveal about the nature of metaphor. In Alan Cienki & Cornelia Müller (eds.), *Metaphor and gesture*, 219–245. Amsterdam: John Benjamins.

Nuñez, Rafael & Eve E. Sweetser. 2006. With the future behind them: Convergent evidence from Aymara language and gesture in the crosslinguistic comparison of spatial construals of time. *Cognitive Science* 30(5). 401–450. doi:10.1207/s15516709cog0000_62

Özyürek, Asli. 2002. Do speakers design their cospeech gestures for their addressees? The effects of addressee location on representational gestures. *Journal of Memory & Language* 46. 688–704. doi:10.1006/jmla.2001.2826

Parrill, Fey. 2007. Metagesture: An analysis of theoretical discourse about multimodal language. In Susan D. Duncan, Justine Cassell & Elena T. Levy (eds.), *Gesture and the dynamic dimension of language*, 83–89. Amsterdam: John Benjamins. doi:10.1075/gs.1.09par

Parrill, Fey & Eve E. Sweetser. 2004. What we mean by meaning: Conceptual integration in gesture analysis and transcription. *Gesture* 4(2). 197–219. doi:10.1075/gest.4.2.05par
Perniss, Pamela & Asli Özyürek. 2015. Visible cohesion: A comparison of reference tracking in sign, speech, and co-speech gesture. *Topics in Cognitive Science* 7. 36–60. doi:10.1111/tops.12122
Ping, Raedy & Susan Goldin-Meadow. 2010. Gesturing saves cognitive resources when talking about nonpresent objects. *Cognitive Science* 34. 602–619. doi:10.1111/j.1551-6709.2010.01102.x
Priesters, Matthias. 2013. *Functional patterns in gesture space*. Institut für Sprach- und Kommunikationswissenschaft der RWTH Aachen. Thesis.
So, Wing Chee, Sotaro Kita & Susan Goldin-Meadow. 2009. Using the hands to identify who does what to whom: Gesture and speech go hand-in-hand. *Cognitive Science* 33. 115–125. doi:10.1111/j.1551-6709.2008.01006.x
Stec, Kashmiri & Mike Huiskes. 2014. Co-constructing referential space in multimodal narratives. *Cognitive Semiotics* 7(1). 31–59.
Sweetser, Eve E. 1992. English metaphors for language: Motivations, conventions, and creativity. *Poetics Today* 13(4). 705–724. doi:10.2307/1773295
Sweetser, Eve E. 1998. Regular metaphoricity in gesture: Bodily-based models of speech interaction. *Actes du 16 Congres International des Linguistes* 5.
van Gelder, Tim. 1998. The dynamical hypothesis in cognitive science. *Behavioral and Brain Sciences* 21. 615–665. doi:10.1017/S0140525X98001733

## Appendix

**Table A1.** ELAN tiers and controlled vocabulary

| Tier | Controlled vocabulary |
|---|---|
| Speech | No controlled vocabulary |
| Gesture type | Iop: iconic observer viewpoint path for path<br>Ios: iconic observer viewpoint path for shape<br>Ioss: iconic observer viewpoint shape for shape<br>Ic: iconic character viewpoint<br>D: deictic<br>B: beat<br>M: metaphoric iconic<br>E: emblem |
| Gesture location | L: left<br>R: right<br>C: center<br>LR: occupying both left and right space (only possible for two handed gestures) |
| Gesture hand | l (left), r (right), b (both) |

| Tier | Controlled vocabulary |
|---|---|
| Gesture referent | Coded based on co-occurring speech as:<br>H: hearing<br>S: sight<br>P: preferred<br>G: good<br>B: bad<br>C: cycle or sequence of related events<br>S: agree<br>D: disagree<br>O: other |

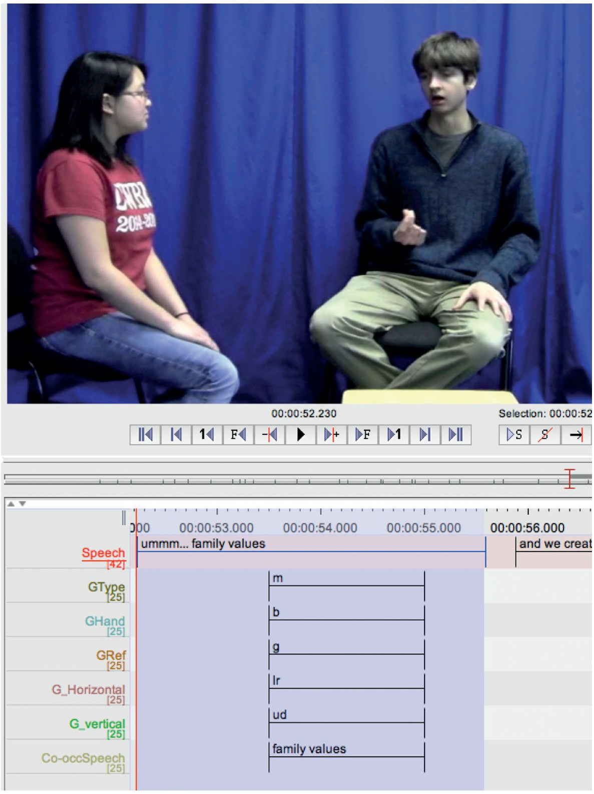



**Figure A1.** ELAN screenshot

## Calculation for measure of overall consistency

Table A2 shows the distribution of gesture referents according to hand used for one speaker, in order to illustrate the overall measure of consistency reported as our second analysis. For *bad things*, this speaker used both hands 3/4 times, so both hands would be used as the consistent choice for that referent. For *cycle*, the consistent choice could be both hands or right hand – the selection doesn't matter for the analysis. For *good things*, the speaker uses both hands 13 out of 23 times, so both hands would be selected, and so forth for the rest of the referents. The proportion row is created by dividing the largest number by the total number of gestures for that referent. These proportions are then averaged to create each participant's mean tendency towards consistency. We can then compare this tendency to expected frequencies. Table A3 shows the expected frequency calculation for the same participant. Expected frequencies are calculated by multiplying row totals (hands used) by column totals (gesture referents), then dividing by total number of gestures for that participant. The same proportion can be calculated (largest number in that column divided by column total), then averaged over referents for a participant. This measure is an average expected frequency for each participant. If participants tend to use the same hand for the same referent, their actual frequency should be higher than expected frequency. We used a paired t-test to compare these two numbers across participants, so a paired t-test would then compare .68 (actual) to .63 (expected) for this participant. Why a paired t-test? A Chi-squared or Fischer Exact test might seem more appropriate, as those tests are designed to explore observed versus expected frequencies. Such an analysis would compare each participant's actual and expected frequencies, but would yield a series of participant-specific measures of consistency, rather than an overall measure. An overall measure was our goal with this analysis, hence the use of a paired t-test. The location analysis used the same procedures.

**Table A2.** Actual frequencies for example participant

| **Hand** | **Bad** | **Cycle** | **Good things** | **Hearing** | **Other** | **Sight** | **Total** | |
|---|---|---|---|---|---|---|---|---|
| Both | **3** | **1** | **13** | **15** | **2** | **6** | | 40 |
| Left | 0 | 0 | 3 | 0 | 0 | 0 | | 3 |
| Right | 1 | 1 | 7 | 5 | 0 | 6 | | 20 |
| **Total** | **4** | **2** | **23** | **20** | **2** | **12** | | **63** |
| Prop. | .75 | .5 | .57 | .75 | 1 | .5 | Mean frequency: | .68 |

**Table A3.** Expected frequencies for example participant

| **Hand** | **Bad** | **Cycle** | **Good things** | **Hearing** | **Other** | **Sight** | **Total** | |
|---|---|---|---|---|---|---|---|---|
| Both | **2.54** | **1.27** | **14.60** | **12.70** | **1.27** | **7.62** | | 40 |
| Left | .19 | .10 | 1.10 | .95 | .10 | .57 | | 3 |
| Right | 1.27 | .64 | 7.30 | 6.35 | .63 | 3.81 | | 20 |
| **Total** | **4** | **2** | **23** | **20** | **2** | **12** | | **63** |
| Prop. | .63 | .63 | .63 | .63 | .63 | .63 | Mean frequency: | .63 |

## *Author's address*

Fey Parrill
Associate Professor
Case Western Reserve University
10900 Euclid Heights
Cleveland, 44106
United States
fey.parrill@case.edu

## *Biographical notes*

**Fey Parrill** is an associate professor of Cognitive Science at Case Western Reserve University. Her work focuses on co-speech gesture within the framework of embodied cognition.

**Kashmiri Stec** defended her PhD on multimodal viewpoint (cum laude) at the University of Groningen in 2016 and now works on early concept development as a UX Researcher at Bang & Olufsen.