Dual viewpoint gestures

Fey Parrill
Case Western Reserve University

This paper examines gestures that simultaneously express multiple physical perspectives, known as dual viewpoint gestures. These gestures were first discussed in McNeill’s 1992 book, *Hand and mind*. We examine a corpus of approximately fifteen hours of narrative data, and use these data to extend McNeill’s observations about the different possibilities for combining viewpoints. We also show that a phenomenon thought to be present only in the narrations of children is present in the narrations of adults. We discuss the significance of these gestures for theories of speech-gesture integration.

**Keywords:** gesture, viewpoint

Dual viewpoint gestures

This paper examines hand and body gestures that simultaneously express multiple perspectives on an event or scene. These gestures, known as dual viewpoint gestures, suggest that a speaker is taking multiple spatial perspectives on a scene at the same time. Despite the fact that this is a rather impressive cognitive feat, relatively little has been written about the phenomenon (though see McClave, 2000). In an effort to provide systematic data on dual viewpoint gestures, we examine a corpus of approximately fifteen hours of narrative data (containing over four thousand gestures). We use these data to extend previous descriptions of the ways in which viewpoint can be combined in gesture. We also show that a phenomenon thought to be present only in the narrations of children is in fact present in the narrations of adults. The paper is organized as follows. We first explain what it means for gesture to express viewpoint, commenting on different uses of the terms viewpoint and perspective. We then describe dual viewpoint gestures, and provide examples from our corpus. We end with some open questions that may serve as starting points for future research.
**Viewpoint in gesture**

Viewpoint, by which we mean the locus of consciousness for model of the world (a definition closely related to that of Chafe, 1994), is ubiquitous in gesture. Because the speaker’s body typically serves as a default deictic center (Bühler, 1982), many gestures encode viewpoint whether or not it is actively part of the speaker’s conceptualization. For example, a speaker who gestures away from her body while describing the path an object took may not have said anything about the object’s point of origin, but a point of origin is nonetheless implicit in her gesture. In addition to gestures performed relative to a speaker-centered origo, researchers have long been aware of the existence of a basic alternation between a more external, third-person representation and a more internal, first-person representation. This alternation has been observed in descriptions of space and in the use of diagrams and models (Bryant & Tversky, 1999). In giving a route description, for instance, a speaker may describe the route by tracing it with a pointed finger (external perspective), or she may orient her body as though moving through the space she is describing (internal perspective). The same basic alternation occurs in narrative language as well. In narrative language, external gestures have been called *observer viewpoint gestures* (O-VPT), while internal gestures have been called *character viewpoint gestures* (C-VPT) (McNeill, 1992). With observer viewpoint gestures, narrators depict an action as though observing it from afar, e.g., by tracing a path. With character viewpoint gestures, on the other hand, narrators use their own bodies in depicting an event. Figures 1b and 1c show this contrast for gestures.
produced in describing the cartoon still shown in Figure 1a. In Figure 1b, the narrator traces the skunk's path with his finger. In Figure 1c, the speaker depicts the action as though he were the skunk.

Figure 1b. Observer viewpoint gesture depicting hopping

Figure 1c. Character viewpoint gesture
In discussing the alternation between an internal and external representation, researchers have used different terminologies. Table 1 provides a representative summary of terms employed in describing the basic internal / external alternation in speech, gesture, and American Sign Language. (Of course many other terms exist for describing spatial frames of reference: see Kita, 2003; Levinson 1997). While the general tendency is to use viewpoint and perspective more or less interchangeably, McNeill distinguishes between the two. Further discussion of this point is warranted, as our use of viewpoint is somewhere between the two.

For McNeill, viewpoint is “the feeling of distance from the narrative” (1992, p. 118), while perspective is “where the observer stands” (1992, p. 192). Viewpoint (as expressed in gesture) can be character’s or observer’s, and is diagnosed based on whether the narrator’s body is incorporated into the representation. Perspective can be either outside (the observer is outside the scene) or inside (the observer

Table 1. Terms used in describing viewpoint and perspective in speech, gesture, or ASL

<table>
<thead>
<tr>
<th>Source</th>
<th>Definition</th>
<th>External Term</th>
<th>Internal Term</th>
</tr>
</thead>
<tbody>
<tr>
<td>McNeill, 1992</td>
<td>Perspective: where observer stands</td>
<td>Outside: transverse motion, body is not origo. Can apply to O-VPT gesture, but not C-VPT gesture</td>
<td>Inside: sagittal motion, body is origo. Always true of C-VPT, can be true of O-VPT</td>
</tr>
<tr>
<td>Kita &amp; Özyürek, 2003</td>
<td>Perspective: no definition given</td>
<td>Event external: from perspective of viewer of stimulus (transverse motion)</td>
<td>Event internal: from perspective of protagonist / actor (sagittal motion)</td>
</tr>
<tr>
<td>Emmorey, Tversky, &amp; Taylor, 2000</td>
<td>Perspective: no definition given, viewpoint also used</td>
<td>Survey: viewpoint is outside scene</td>
<td>Route: viewpoint is within scene</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Diagrammatic space: ASL signer is outside the space — correlates with survey, authors note similarity to O-VPT</td>
<td>Viewer space: ASL signer is inside the space, above — correlates with route, authors note similarity to C-VPT</td>
</tr>
<tr>
<td>Bryant &amp; Tversky, 1999</td>
<td>Perspective: no definition given, viewpoint also used</td>
<td>Outside: intrinsic, narrator not part of scene</td>
<td>Inside: egocentric, narrator part of scene</td>
</tr>
<tr>
<td>McNeill, 1992</td>
<td>Viewpoint: feeling of distance from the narrative</td>
<td>O-VPT: narrator’s body is excluded, narrator is outside story (can be inside or outside perspective)</td>
<td>C-VPT: narrator’s body is incorporated, narrator is inside story (must be inside perspective)</td>
</tr>
</tbody>
</table>
is within the scene). If the narrator produces a C-VPT gesture, the perspective is inside the scene by default. If a narrator produces an O-VPT gesture, the perspective is likely to be outside the scene, but may also be inside the scene. For example, if a narrator traces the trajectory of an object starting at her shoulder and moving away from her body, her body has not been incorporated into the scene so the gesture is O-VPT. However, the frame of reference imposed by her body has been exploited (the direction of motion is relative to the body), so the perspective is inside.

We are interested in the locus of consciousness for a model of the world that the narrator is representing, which will be either inside the story world or outside it. We will thus assume C-VPT gestures indicate that the narrator is inside the story world, while O-VPT gestures indicate that she is outside the story world.

Character and observer viewpoint gestures are extremely common. Less common, however, are gestures in which a narrator combines multiple perspectives, known as dual viewpoint gestures (McNeill, 1992). While these gestures are rare, they illustrate a number of interesting possibilities for combining character and observer viewpoint.

**McNeill’s dual viewpoint types**

McNeill (1992) identifies two basic types of dual viewpoint gestures. The first type combines some aspect of a character’s viewpoint with the same character’s trajectory, but depicted from an observer’s point of view. These will be referred to as character + character’s trajectory gestures. In McNeill’s example, a narrator is describing a character’s downward fall. The character is clutching something in his hand as he falls. The narrator combines a grasping handshape (C-VPT) with a trajectory (O-VPT) (p. 123). In the second type, the viewpoint of two different characters is combined. McNeill describes these as chimera dual viewpoint gestures. He provides two instances. The first is from an adult (originally presented in McClave, 1991, and more recently in McClave, 2000). The speaker points at his own body, using the pointing hand to represent one character, while the body represents another character. The speaker says *you had your doctor go over to check out that person’s claim*. The body thus represents the viewpoint of “that person”, while the pointing hand represents the viewpoint of the doctor. The second instance comes from a child, who enacts one character hitting another over the head with an umbrella. However, rather than hitting empty space in front of her body, she enacts hitting herself in the head, thus using her hand to show one viewpoint (the hitter) and her body to show another (the hit-ee).
Our extensions

In this paper, we make three new contributions. First, we examine a corpus of over four thousand gestures to assess the relative frequencies of character, observer, and dual viewpoint gestures. This corpus represents only one discourse genre (narrative) and is therefore hardly comprehensive, but such basic facts about gestural viewpoint have yet to be established, and we provide a beginning. Second, McNeill suggests that when adults produce chimera gestures, they exclusively involve pointing gestures. We will provide an example of a chimera gesture of enactment (akin to the child’s example above) from an adult narrator. Third, while McNeill has focused on gestures that involve a single hand encoding two points of view, or a single hand acting on the body, we suggest a new category of gesture for inclusion. Specifically, multiple gestures produced simultaneously, but with different articulators (e.g., hands and body, hands and legs), also offer possibilities for combining different viewpoints, and should be considered dual viewpoint gestures as well.

This point requires some clarification. The handshape + trajectory gesture described above (in which the character falls while holding something) can be seen as involving one articulator, the hand, in which two points of view are present. But the nature of character viewpoint means that the narrator’s body has “become” the character’s body. This gesture therefore can also be seen as involving two articulators, the hand (which shows the path), and the body, which represents the character’s body. Once dual viewpoint is seen in this light, there seems to be no reason to exclude gestures in which the body takes on one point of view and the hand another, or gestures in which the body and hand illustrate different aspects of the same point of view. We will refer to such gestures as dual viewpoint, multiple viewpoint.

<table>
<thead>
<tr>
<th>Type</th>
<th>Example</th>
</tr>
</thead>
<tbody>
<tr>
<td>One articulator (single gesture)</td>
<td></td>
</tr>
<tr>
<td>O-VPT + C-VPT: Character + same character’s trajectory</td>
<td>See McNeill, 1992, pp. 123</td>
</tr>
<tr>
<td>O-VPT + C-VPT: Character + another character’s trajectory</td>
<td>Not observed</td>
</tr>
<tr>
<td>C-VPT + C-VPT: Chimera with point</td>
<td>See McNeill, 1992, pp. 124</td>
</tr>
<tr>
<td>C-VPT + C-VPT: Chimera in children</td>
<td>See McNeill, 1992, pp. 318</td>
</tr>
<tr>
<td>C-VPT + C-VPT: Chimera in adults</td>
<td>See Example 1</td>
</tr>
<tr>
<td>Two articulators (simultaneous gestures)</td>
<td></td>
</tr>
<tr>
<td>O-VPT + C-VPT: Character + same character’s trajectory</td>
<td>See Example 2</td>
</tr>
<tr>
<td>O-VPT + C-VPT: Same character, trajectory + manner decomposition</td>
<td>See Examples 3 and 4</td>
</tr>
<tr>
<td>O-VPT + C-VPT: Character + another character’s trajectory</td>
<td>See Examples 5 and 6</td>
</tr>
</tbody>
</table>
articulator gestures (as distinguished from dual viewpoint, single articulator gestures). There appear to be several kinds of such gestures.

The first is a character + character’s trajectory gesture, where viewpoint is split across hand and body. This phenomenon has a subtype, where different aspects of a single character’s action are encoded by different articulators. The second is the character + another character’s trajectory gesture, where two points of view are split across hand and body. In Table 2, we summarize the possibilities for combining multiple viewpoints.

What dual viewpoint isn’t

In the next section, we describe our corpus and illustrate those possibilities for combining viewpoint not already discussed in McNeill, 1992. Before going further, however, we should comment on several phenomena that are not included in our categories of dual viewpoint gestures. First, we chose to focus on dual viewpoint gestures where the hands and body are the loci of viewpoint. That is, we chose to exclude facial gestures (including gaze phenomena). This is primarily because, while coding schemes for hand and body gestures are well established, coding schemes for facial gesture and for the relationship between facial and hand/body gestures are not.

Second, although McNeill does not make this point explicit, dual viewpoint gestures involve entities that can potentially have a point of view. That is, not every gesture that involves two entities is considered a dual viewpoint gesture, even if

Figure 2. Character (right hand) observer (left hand) non-dual viewpoint gesture: only one entity with point of view
one entity is depicted from an observer’s point of view and one from a character’s. For example, a narrator who describes a character catching a ball may use character viewpoint to show the character’s hand, but observer viewpoint to show the ball, as shown in Figure 2. Because the ball is not an entity that can have a point of view, however, such gestures are excluded. The only point of view represented is that of the narrator-as-character.

Finally, two separate characters are sometimes represented from an observer’s point of view. These gestures involve only one point of view (that of an observer viewing the event from a distance), and are therefore not considered dual viewpoint gestures.²

Corpus and examples

Our corpus contains 15 hours of narrative descriptions of cartoons. Data come from 27 University of Chicago students and 124 Case Western Reserve University students. Participants watched cartoon clips or read cartoon descriptions, then described these stimuli to a friend. All participants were native speakers of English, and all were paid for participating.

Data analysis

Participants were video/audio recorded as they described the stimuli to their partners. Gestures in which participants took on the role of the character, using their bodies as the character’s body, were coded as character viewpoint gestures (C-VPT). Gestures in which the participant took on the role of an observer, usually with hand or arm movements in the space in front of the body to reflect the character as a whole, were coded as observer viewpoint gestures (O-VPT). Gestures in which both character and observer viewpoint were present (regardless of the number of articulators involved) were coded as dual viewpoint gestures (D-VPT). Gestures with no motion event content (rhythmic beat gestures, metaphoric gestures, deictic gestures, or iconic gestures that simply traced shapes) were classified as no viewpoint gestures (N-VPT). The data were classified in this manner by two independent coders. Coders agreed on 78% of all gestures, and disagreements were resolved through discussion.

Summary data

We first provide summary data on the relative occurrence of all gesture types. Table 3 shows both the raw frequencies of each type, and the mean proportion
of each type (that is, the proportion of each type appearing in a participant’s data, averaged across all participants). For these stimuli, observer viewpoint gestures and character viewpoint gestures appear with roughly equal frequency, while dual viewpoint gestures are extremely rare. Table 4 shows the distribution of the different dual viewpoint types described above.

**Examples of dual viewpoint gestures**

**Example 1: Single articulator**

The following example illustrates a chimera (dual C-VPT) gesture with enactment in an adult. The narrator is describing a mouse zipping a cat’s mouth shut (Figure 3a). Her right hand (Figure 3b) enacts the event with reference to her own mouth.

Speech: *[zipped its mouth shut]*

Gesture: Right hand moves across mouth
Example 2: Multiple articulators: character + character’s trajectory
In this example, the body also moves, rather than simply serving as a reference point. The narrator is describing a character’s trajectory (Figure 4c), and does a very exaggerated body lean to mimic the action of the character (C-VPT). At the same moment, his hand traces the path taken by the character (O-VPT). The preparation (Figure 4a) and the end of the stroke phase of the gesture (Figure 4b) are shown to make the degree of movement apparent.
Speech: *the cat [you know *banks a turn] on the wall*
Gesture: Body leans significantly to left while right hand traces cat’s trajectory

*Examples 3 and 4: Multiple gestures: trajectory and manner decomposition*
In the next two examples, the same character’s action is encoded from multiple points of view, but in this case, the action is decomposed into separate elements. The first element is the trajectory, and the second is the internal structure of the...
motion. In his analysis of motion event components, Talmy (1985) has referred to these as path and manner, respectively.

There is not necessarily a clear distinction between these examples and the body lean above, which could be viewed as manner of motion. However, the following two instances are classic cases of manner, so have been treated separately. In the scene described in Example 3 (Figure 5a), a character moves down a street with his legs gyrating. The speaker encodes character viewpoint and manner of motion by pumping his legs, and encodes observer viewpoint by tracing the trajectory (Figure 5b). In other words, he encodes different aspects of the same complex motion event using different articulators. While combinations of manner and path within a single gesture are not unusual in describing this event (Duncan, 1996; Kita & Özyürek, 2003; Kita, Özyürek, Allen, Furman, & Brown, 2005; McNeill, 1992; McNeill & Duncan, 2000; Parrill, 2008), decomposition across multiple articulators is rare, and the legs are hardly ever used as articulators.

Example 3 Speech: down the street...[through*] across the street
Gesture: Right hand traces trajectory, legs pump up and down as though pedaling.
In Example 4, the narrator has read a cartoon description text, and is describing the following event from the text: *When he gets out of the stands he knocks over a hot dog cart and flails around as he tries not to trip on all the hotdogs.* The narrator produces a small path gesture with her right hand while her left hand moves back
and forth, shown in Figure 6. (The wavy path she traces with the right hand also encodes manner of motion but, for the sake of simplicity, we focus on the trajectory aspect of this gesture.)

Example 4 Speech: *he finally like* [slides his way out]

Gesture: Right hand shows trajectory of rabbit sliding, while left is character’s hand flailing as he attempts to keep his balance.

![Figure 6. Trajectory (right hand) + manner (left hand) decomposition](image)

**Examples 5 and 6: Multiple gestures, character + another character's trajectory**

The following two examples are virtually identical. The narrators are describing one character escaping from another’s arms (Figure 7c). One arm is used in a stroke hold (a stroke where the effortful part of the gesture is static: McNeill, 2005) character viewpoint gesture, while the other traces a different character’s path, but from an observer’s viewpoint.

Example 5 Speech: *she slips out of his arms* [pshoo]

Example 5 Gesture: (repeated twice) Left hand encodes skunk cradling cat, while right shows trajectory of cat leaving skunk’s arms (Figures 7b and 7c).

Example 6 Speech: *scurries like* [out of his grasp]

Example 6 Gesture: See above.
Figure 7a. Cartoon event: cat escapes from skunk

Figure 7b. Character (left hand, body) + another's trajectory (right hand)
Discussion

This study has been primarily descriptive, thus leaves a great many questions unanswered. The first might be: What circumstances elicit dual viewpoint gestures? These data don’t allow us to provide much of an answer. Our corpus contains many descriptions of each particular event described above, yet dual viewpoint gestures are rarely performed for any specific event. While events in which a character acts on the head or face of another character (as in Example 1) may be more likely to evoke chimera gestures, we have no explanation for the fact that only one narrator use her own body when talking about the event in Example 1, while all the others use the space in front of their bodies. We have also shown that a phenomenon suggested to be unique to children — dual C-VPT gestures that do not involve a deictic point — also occurs in adults. Given the rarity of dual viewpoint gestures in general, we suspect that examples of such gestures in adults simply did not occur in the dataset on which McNeill based this suggestion.

One might also ask why these gestures are so infrequent. Given the option of encoding a great deal of information simultaneously, why do people choose to present more schematic versions of events? We do know from previous research that speakers often distribute information across the manual and vocal modalities (Duncan, 1996; Kita & Özyürek, 2003; Kita, Özyürek, Allen, Furman, & Brown, 2005; McNeill & Duncan, 2000; Parrill, 2008). This tendency might help to explain why dual viewpoint gestures are not very common. A related question might be:
Are these gestures more cognitively demanding? Is the load on working memory or on attentional resources greater for speakers when they are combining multiple viewpoints (and specifically multiple character’s viewpoints)? The fact that verbal dysfluency accompanies one of our examples suggests that this might be the case, but is far from conclusive. This question could potentially be tested using established cognitive load paradigms (Goldin-Meadow, Nusbaum, Kelly, & Wagner, 2001; Wagner, Nusbaum, & Goldin-Meadow, 2004). Similarly, we might ask whether these gestures are more informative for speakers. Beattie and Shovelton (Beattie & Shovelton, 2001, 2002) have shown that character viewpoint conveys more information to viewers: Is this also the case with dual viewpoint gestures?

Another avenue of research might involve considering how such gestural phenomena compare to so-called body partitioning (Dudis, 2004) in American Sign Language. With body partitioning, conventional articulators (face, hands) can be partitioned off to simultaneously represent information from multiple perspectives. For example, the face might take on character viewpoint while the hands use lexical devices (e.g., so called classifiers) to show an observer’s point of view. Given that these behaviors are usually linguistic (rather than being gestural), we might ask how entry into a linguistic system changes the possibilities for combining points of view.

There are also interesting analogues in speech, in which two different representations seem to be simultaneously activated. Such situations have been referred to as syntactic blends (Bolinger, 1987; Cutler, 1982; Fay, 1981; Tuggy, 1996). A classic example is the double copula construction: a sentence like the problem is that I’m hungry is subject to a certain analogical pressure from a sentence like what the problem is, is that I’m hungry, resulting in the problem is is that I’m hungry when both constructions are activated at once. Such cases are not a phenomenon of viewpoint, as defined here, but suggest that a similar process can occur during speech production. Dual viewpoint gestures might therefore be profitably analyzed as a kind of gestural blend.

We have described in detail a phenomenon that is relatively rare, and a reasonable reader might wonder whether such an illusive creature as the dual viewpoint gesture truly merits this attention. We suggest that just as the study of idioms can inform our understanding of compositionality in language, the study of rare combinations of semiotic features in gesture can inform our theories of gesture-speech integration.
Notes


2. In our corpus these gestures are also quite rare (approximately .04%).

3. Gesture occurs during the bracketed speech, peak prosodic emphasis is in bold (see McNeill, 1992).

4. I thank an anonymous reviewer for suggesting this parallel.

References


Tuggy, David (1996). The thing is is that people talk that way. The question is is why? In Eugene H. Casad (Ed.), *Cognitive linguistics research 6: Cognitive linguistics in the redwoods* (pp. 714–752). Berlin: Mouton.


**Author’s address**

Fey Parrill  
Department of Cognitive Science  
Case Western Reserve University  
10900 Euclid Ave.  
Cleveland, OH 44106  
USA  
fey.parrill@case.edu

**About the author**

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