

Meaningful shifts

A review of viewpoint markers in co-speech gesture and sign language

Kashmiri Stec

University of Groningen

This review describes the primary strategies used to express changes in conceptual viewpoint (Parrill, 2012) in co-speech gesture and sign language. We describe the use of the face, eye gaze, body orientation and hands to represent these differences in viewpoint, focusing particularly on McNeill's (1992) division of iconic gestures into observer versus character viewpoint gestures, and on the situations in which they occur. We also draw a parallel between the strategies used in co-speech gesture and those used in different signed languages (see Cormier, Quinto-Pozos, Sevcikova, & Schembri, 2012), and suggest possibilities for further research in this area.

Keywords: gesture, sign language, viewpoint

Imagine you are walking outside on a sunny day. Someone is skydiving, intending to land on a grassy field nearby. You tilt your head up to watch his descent, and see his parachute deploy. You watch him skillfully navigate the currents of air and then, suddenly, narrowly miss colliding with an unidentified object. Despite the scare, he completes his descent and lands safely in the field. You breathe a sigh of relief and continue your walk. Later, you tell this story to a friend. How do you do it? How do you describe each participant's actions and reactions to these events? Do you relate your own perspective, as witness to them, or do you rather relate what the skydiver did as he made the narrow escape and safe landing?

These questions concern the conceptualization of events, particularly regarding the *viewpoint*, or physical perspective, from which a conceptualizer represents a scene, as well as the means by which that viewpoint is expressed. Following Sweetser (2012), *viewpoint* may also refer to the general phenomena of affordance access granted by one's physical location: One's physical location isn't only visual

information, but also implies what one can reach and interact with from that location. Along these lines, recent experimental work in cognitive science (e.g., Bergen, 2012) has shown that anybody processing language does so from a particular embodied perspective which typically takes one of three forms: a first-person perspective, a third-person perspective, or a combination of different perspectives.

Considering the complex array of viewpoints and viewpoint choices available to us (see, e.g., Chafe, 1994, or Parill, 2012, for a review), one question we can ask is: How are these different viewpoints represented in face-to-face communication, and by what means? Another is: what do these differences in viewpoint afford us (in terms of learning, clarity, information content, etc.)?

To begin to answer these questions, we can look to the ways viewpoint is represented in co-speech gestures and sign languages. Of interest here are the cues provided by visible bodily action which allow speakers to gesture as if they are taking one particular viewpoint or another, as well as how these cues interact with lexical expressions, whether spoken or signed, to produce multi-modal utterances which are somehow “consistent” in their depiction of viewpoint. Just as spoken language provides a variety of means by which viewpoint can be expressed, e.g., as if the speaker or signer were a participant or observer of a particular event, so too can visible bodily actions convey these differences in viewpoint. Using a variety of visible bodily actions, one can enact the body of a participant or observer of a particular event. These actions range from facial displays to variations in gaze, body orientation and iconic gestures, as well as their interaction with conceptual structures such as frames of reference. Together with lexical expressions, these comprise the viewpoint of the multi-modal utterance.

For example, in the skydiving scenario described above, gestural viewpoint could be shown by moving your arms around as if fiddling with parachute controls with a look of terror on your face (first-person, skydiver’s perspective); tracing the skydiver’s path as he descends, swerves and lands (third-person, your perspective); or showing surprise on your face as you trace the skydiver’s path through the air (mixed).

Looking at the encoding of viewpoint of multi-modal utterances in detail, Parrill (2012) describes *viewpoint* as consisting of three complementary aspects: conceptual viewpoint, linguistic viewpoint and gestural viewpoint. *Conceptual viewpoint* is the real or imagined physical location of a conceptualizer; it denotes the mental simulation of an event, including the mental images, motor program and representations of mental states which make that simulation possible. Often, these are made with specific physical locations in mind. *Linguistic viewpoint* is indicated by the various linguistic devices (lexical choice, tense, etc.) which suggest the conceptual viewpoint adopted by the conceptualizer. Linguistic viewpoint coincides with conceptual viewpoint insofar as specific physical locations lead to

different images, and thus, different descriptions based on them (e.g. as a participant or observer of an event). *Gestural viewpoint* is indicated by the use of the body to encode different perspectives. Like linguistic viewpoint, gestural viewpoint may suggest the conceptual viewpoint adopted by the conceptualizer, e.g., as a participant or observer of an event, based on the visible bodily actions which are used.

The visual perspective denoted by conceptual viewpoint may be highly decontextualized (consider generic third-person narrators), and may be suggested in language by any number of means (for a review, see Chafe, 1994; Earis, 2008; or Parrill, 2012). These linguistic cues can influence the comprehension of a given sentence or scene (Bergen, 2012), especially at significant narrative moments (Bundgaard, 2010). It stands to reason that gestural viewpoint, which iconically represents the imagined body of a participant or observer of a particular event, should do the same.

Recent work (Hostetter & Alibali, 2008) describes co-speech gestures as a feature of the motoric and perceptual simulation systems which underlie language and embodied thought (see Barsalou, 1999). They are often used by speakers to highlight visual and spatial features of the information provided in words. As Enfield, Kita, & de Ruiter (2007) suggest, the degree to which a given piece of information is in focus may affect various properties of the gesture, such as its location and size, and thus may promote a focus on the simulation of a body and its particular viewpoint, located somewhere in space.¹ As Sweetser (2012) points out, not only are we constantly aware of own physical viewpoint, most of us are also aware of the physical viewpoints of those around us. Hence, depicting viewpoint is not only a subjective activity, but an inherently intersubjective one (see Verhagen, 2005).

In light of this, we consider the relationship between conceptual and gestural viewpoints, and discuss the strategies used in co-speech gesture and sign language to express those viewpoints, providing both descriptions and experimental evidence of their use. In the case of co-speech gesture, we also discuss several pedagogical findings which relate multi-modal viewpoint expression to depth of understanding. We finish with a comparison of the similarities and differences in strategies used in co-speech gesture and sign language, and suggest areas for future inquiry.

Viewpoint strategies in co-speech gesture

Several categories of manual gestures have been proposed (see Kendon, 2004, for an historical survey of different approaches to gesture categorization). Here we follow McNeill's (1992) division which distinguishes emblems, beats, points and

other deictic markers from iconics. In this section, the use of iconic gestures and other visible bodily actions are considered which allow the expression of a participant or observer of an event. These visible bodily actions have iconic properties (see McNeill, 1992; Mittelberg, 2006; Parrill & Sweetser, 2004; Sweetser, 2006) which readily link the form of the gesture to its referent, such as a speaker who says “turn the pages of the book” and gestures by repeatedly moving their left hand in an arc from their right to left, the usual motion made when reading a book in the West. These visible bodily actions also tend to evoke meaning that is global, image-based, and dependent on context to a greater extent than is language (Kendon, 1980; McNeill, 2005).

Because speech and co-speech gestures can often express the same underlying idea without necessarily expressing identical aspects of it (Hostetter & Alibali, 2008; McNeill & Duncan, 2000), there need not always be a one-to-one mapping between linguistic and gestural viewpoint. Although as analysts we might tease apart these mechanisms in order to understand how viewpoint is expressed and maintained, we should remember that dialogue is “opportunistic”, meaning that interlocutors “use whatever works in the moment” (Bavelas & Chovil, 2000, p. 188). But even this “opportunistic” use may be structured. In this section, we will consider how viewpoint is expressed in co-speech gesture: the means by which it is expressed, the conditions which affect its expression, and the situations in which differences in how viewpoint is expressed in gesture have been shown to have profound effects on cognition.

Means of expression

Bodies are powerful anchors for conceptual viewpoint (Sweetser, 2012) — the direction of one’s gaze has very real effects on what one sees and what information is accessible from that vantage point. In addition to the direction of gaze, each of the primary articulators of the body may also be used to demonstrate different aspects of viewpoint. In this section we will consider each of the following in turn: gaze and facial expression, several kinds of manual gestures, whole body enactments, and gestures which simultaneously express multiple viewpoints.

Gaze and facial expression

The face, and gaze in particular, are rich sources of viewpoint information. Often, due to their proximity, they work in coordination. Sidnell (2006) notes that, in English, quotatives are often accompanied by facial displays of affect while averted gaze is often used to mark the left boundary of quotes, and gaze returned to the addressee is used to mark the right boundary. He also notes that when speaker gaze is directed towards the narrative space established by the quotative, addressees

change their gaze behavior, too: instead of tracking the speaker's gaze, addressees look to the gestures and the narrative space that are evoked. In Korean, Park (2009) observes differences which depend on the source of the quote: when quoting one's past self, speakers avert their gaze and use overly expressive facial expressions as a sort of caricature, but when quoting a co-present person's past speech, speakers use exaggerated prosody, touch the person they are quoting as if to ground or anchor the quote, and turn their gaze to the addressee(s). At the same time, the person who has been quoted removes their gaze from the speaker and also turns to gaze at the addressee(s). Along these lines, both Maury-Rouan (2011) and McClave (2000) note that explicit and implicit expressions of viewpoint shift are accompanied by head tilts to the side while the face is re-oriented. Maury-Rouan (2011) further notes that gaze is removed only in explicit viewpoint shifts, such as when quoting; otherwise, it remains fixed on the addressee.

Turning to facial displays, Bavelas & Chovil (2000) note that for English speakers, facial displays tend to be partial portrayals insofar as certain stylized features, such as raised brows, are used to suggest emotions like surprise rather than express them directly. In this way, facial displays tend to be used to evoke a sense of emotion or state of mind in the same way that character viewpoint gestures are used to evoke a sense of a character's body or movements. For example, if in the middle of a narrative about a car crash which was narrowly avoided, a speaker says, "We're going 'Oh my god!'" while making a fearful face — rounded mouth, wide eyes, raised brows, gaze directed away from the addressee — that face refers to events in the narrative, not to the act of retelling (Bavelas & Gerwing, 2007, p. 299). Bavelas & Chovil (2000) classify facial expressions into two groups: iconic facial displays, which depict the speaker at another point in time or another person, and facial metaphors, which are a sort of meta-comment offered by the speaker, e.g., when squinted eyes and furrowed brows are used to communicate incredulity. Importantly, these facial displays are related to the structure of the talk and do not reflect the emotional state of the quoted speaker at the time of the original utterance. In this way, facial expressions and even gaze can be used as a means of expressing differences in conceptual viewpoint.

Manual gestures: character vs. observer viewpoint

Manual gestures may be categorized in different ways. Here, we focus on McNeill's iconic gestures (McNeill, 1992) which are often further subdivided into the following categories: character viewpoint gestures (CVPT) and observer viewpoint gestures (OVPT).

CVPT gestures map the character's body onto the gesturer's body such that the gesturer's hands, torso, gaze and etc. correspond to those of the character (McNeill, 1992). Thus, the gesturer's body parts are the character's body parts (see

Dudis, 2007, or Liddell, 2003, for a discussion on how this is possible). This iconic mapping enables the experience of characters “as they were”: speakers re-enact them from a first person perspective, and addressees witness them from a third person perspective, as they would anybody else. For example, after watching a cartoon cat “hop along”, one gesturer (Parrill, 2009, p. 273) demonstrates the cat’s movements with a CVPT gesture: he raises his wrists and turns his palms down, showing his hands as cat paws, and moves his hands up and down, illustrating the hopping action as the cartoon cat might have experienced it.

Unlike CVPT gestures which enable the gesturer to embody a character, OVPT gestures are schematic (McNeill, 1992). Typically, OVPT gestures condense information into one or two dimensions only — e.g., trajectory, manner, shape, static layout or orientation — and show the scene to the interlocutor as if from afar. For example, if we consider the same cartoon cat who hops along, another gesturer (Parrill, 2009, p. 273) demonstrates the cat’s movements by using his right index finger to trace the trajectory of the cat’s hopping movement: vertical oscillations along a horizontal path.

OVPT gestures tend to be used to profile schematic, spatial-orienting qualities while CVPT gestures profile the embodied character and its experiences and reactions. As the examples described here show, even the same stimulus (a cartoon cat, hopping along) may elicit different gestural viewpoints from different participants who are completing the same task. Thus, gestural viewpoint suggests a difference not only in physical viewpoint but in conceptual viewpoint, presumably highlighting different features according to their saliency with respect to the speaker.

Character enactments/embodiments

Several scholars (Clark & Gerrig, 1990; Fox & Robles, 2010; Park, 2009; Sidnell, 2006; Streeck, 2002) note a recent trend towards analytic awareness of mimetic performance in face-to-face communication, which they describe as a preference for depicting events as they were witnessed. What Clark & Gerrig (1990) and Fox & Robles (2010) call demonstrations, and Sidnell (2006) calls depictions (roughly: speech, co-speech gesture and facial displays of affect), are central to these constructions since they give interlocutors as direct an experience of the original event as our bodies allow. Demonstrations include visible gestures, such as CVPT gestures, and also vocal gestures such as pitch, intonation and accent. This allows addressees to make inferences about the event without explicitly relying on the speaker to make those inferences clear, thus “eras[ing] the witness and obscur[ing] the subjective basis of an account” (Sidnell, 2006, p. 406) so that the addressee feels as if they were actually there, witnessing the event happen in real time.

Multimodal expressions of this kind are especially useful for reproducing emotion, urgency, indecision, sarcasm, disfluencies, prosody and gesture. Whereas

verbal descriptions of these would be cumbersome, demonstrations are easy, fluid and clear, thus supporting the observation that demonstrations in “real time” are both easier and preferred (Goldin-Meadow & McNeill, 1999). Like Parrill (2010b) and Clark and Gerrig (1990), Sidnell observes that some events more easily lend themselves to demonstration than others — those which rely on spatial orientation, body part use or physical viewpoint often elicit more enactment-like behavior than others (Sidnell, 2006).

Although it would be interesting to see the degree to which the different viewpoint strategies described here interact with each other to create these enactments, in particular the role that CVPT gestures may play, at the time of this writing thorough descriptive work on enactments in co-speech gesture has yet to be done.

Gestures with multiple viewpoints

Although single-viewpoint gestures are the most frequent, gestures which simultaneously express multiple viewpoints do exist, especially if we include non-manual behaviors such as facial expression (see Bavelas & Chovil, 2000). These gestures partition the body (in the sense of Dudis, 2004), creating multiple real space blends such that each partitioned section or area expresses a different gestural viewpoint. These dual viewpoint gestures (Parrill, 2009) necessarily involve multiple articulators and are characterized by the simultaneous expression of CVPT and OVPT gestures. They may partition the body vertically, horizontally, and even diagonally, and may recruit any body part: arms and hands; legs and feet; the head and face, etc.

Parrill (2009) describes several such gestures. In one, the gesturer is shown representing a cartoon skunk holding a cat who immediately runs away (Parrill, 2009, p. 285). The gesturer represents the body of the cartoon skunk with a CVPT gesture: his left arm curves to hold a body just as the skunk’s right arm does in the cartoon. At the same time, the gesturer’s right hand performs an OVPT gesture which displays the trajectory of the cat as she runs away: a straight line along the sagittal axis, away from the gesturer’s body. This divides the gesturer’s body such that each articulator corresponds to a different character’s body.

But there are also instances in which dual-viewpoint gestures represent different perspectives on the same body. In another example, Parrill describes a dual-viewpoint gesture made to represent a cartoon cat who recently swallowed a bowling ball and is half-running, half-rolling down a hill (2009, p. 283). In this case, the gesturer makes an OVPT gesture with the right hand to illustrate the cat’s trajectory while making a CVPT gesture with his legs and feet to illustrate the motion the cat’s lower body makes while going down the hill.

If we extend Parrill’s category of dual viewpoint gestures to include other visible actions, such as facial expression or gaze, then we see other interesting

partitions of the body. For example, Bavelas and Chovil (2000, p. 167) describe a gesturer who recounts her brother's uncertainty at an Italian train station: the gesturer's hands make an OVPT gesture which shows the relative locations of different trains while the gesturer's eyes, which widen and then bounce back and forth between the trains, show the brother's uncertainty. A puzzled face could easily express some of the same information as the moving gaze. In this case, the gesturer's hands set the scene from an observer's perspective while the head shows a specific non-redundant reaction to it with CVPT.

This simultaneity and multiplicity of viewpoints expressed in gesture recalls Sweetser (2012) and Young's (2002) observations that we are constantly surrounded by other viewpointed bodies whose physical viewpoints affect our own. Even though we only actually experience one physical viewpoint at a time, the one afforded by our own eyes, we are simultaneously aware of other conceptual viewpoints. One viewpoint is often not enough: multiple viewpoints are needed to accurately recreate a scene and present as much information about it as possible.

Viewpoint preferences

If we take seriously Hostetter and Alibali's (2008) claim about the relation between co-speech gestures and mental simulation, then gestural viewpoint, and all of the semantic effects it entails, actually reflects the way that speakers and addressees imagine, simulate and experience a given scene. But what prompts one viewpoint strategy over another? And what effects on cognition might that choice entail? We will consider each question in turn.

Linguistic preferences

Research conducted on gestural viewpoint preference has focused on the CVPT vs. OVPT distinction, and shows that there are both language-specific and generic conceptualization constraints.

Research on the encoding of event properties suggests that the structure of the language one speaks exerts an influence on how these features are distributed across speech and gesture (e.g., Özyürek, Kita, Allen, Furman, & Brown, 2005) as well as on the gestural viewpoint speakers prefer while describing a given event. For example, while English verbs typically encode both path and manner information in the verb root, co-occurring gestures tend to encode only path information from an observer's perspective. This tendency is so strong that, as Brown (2008) reports, even intermediate learners of English adopt this preferential viewpoint strategy in multi-modal communication. She notes in particular that although Japanese speakers tend to prefer co-occurring CVPT gestures, native Japanese speakers who have relatively low exposure to English (basic to intermediate level)

adopt English's OVPT strategy not only when speaking English, but also when speaking Japanese. By doing so, they demonstrate an influence of second language on native language. This contrasts to earlier research (Haviland, 1996; Nuñez & Sweetser, 2006) which, though not concerned with viewpoint per se, demonstrated an effect of gesture styles typical of the speaker's native language affecting gestures in the second language.

Conceptualization constraints include the suggestion by McNeill (1992) that, in American English, CVPT vs. OVPT gestures might reflect the complexity of the utterance or the imagery associated with it, for example by noting that transitive utterances tend to be accompanied by CVPT gestures while intransitive utterances tend to be accompanied by OVPT utterances. He also noted that events which are causally central to discourse tend to be accompanied by CVPT gestures. Parrill (2010b) extended McNeill's English work by describing several factors which affect the correlation between gestural viewpoint and information structures. These are: discourse structure, event structure, and linguistic structure.

Concerning discourse structure: Parrill's (2010b) data suggest that events which are causally central to discourse (e.g., actions or results) are more likely to be accompanied by co-speech gesture than peripheral events (e.g., character introductions or reactions). However, she found no significant difference between viewpoint strategies accompanying these discourse-central events. In cases where a discourse-central event evoked both CVPT and OVPT gestures, Parrill found other factors, such as the narrator's focus, whether the information was shared or new, and the structure of the narrative determined the viewpoint of the gesture. In particular, shared information was less likely to have CVPT and more likely to be accompanied by OVPT gestures (Parrill, 2010a).

In other cases, the represented event itself may preclude one gestural viewpoint or another. Parrill found that the spatial, motion and imagistic nature of an event can affect its portrayal, as can affordances of the different objects in use. She found that motion along a path or trajectory is highly correlated with OVPT gestures, regardless of transitivity. She also noted that some object affordances do not work with some viewpoints, e.g., some objects (like newspapers) and events (like reading them) tend to be more often represented gesturally with CVPT gestures. In fact, in general, if an object or an event involves the upper body or torso, or if the description involves a display of affect, then the accompanying gestures often have CVPT.

Finally, Parrill (2010b) found that linguistic structure also influences gesture type. Her findings support McNeill's (1992) observations on the correlation between gestural viewpoint and transitivity. Looking at a corpus of narrative data based on the *Sylvester* and *Tweety Bird* cartoons, she found that while 56% of transitive utterances included a CVPT gesture, only 33% of intransitive utterances did.

The result is even more striking with intransitive utterances: only 16% were accompanied by a CVPT gesture, while 80% were accompanied by an OVPT gesture. This finding is supported by Beattie and Shovelton (2002), who found a strong correlation between transitivity and gestural viewpoint in British English. In their task, participants first described short cartoons to another participant. Iconic gestures were then extracted (15 each with CVPT or OVPT) from these re-tellings and were played back silently to a second group of participants. This second group of participants then had to describe what they had seen. Interestingly, Beattie and Shovelton found that CVPT gestures were significantly more likely to be accompanied by transitive descriptions than OVPT gestures, which were more likely to be accompanied by intransitive ones. They conclude that the transitivity of the clause is at least partially signaled by the viewpoint of the accompanying co-speech gesture (Beattie & Shovelton, 2002, p. 189).

Cognitive preferences

Concerning the effect of gestural viewpoint on cognitive processes, Beattie and Shovelton (2002) used a memory test to show that CVPT gestures convey more information than OVPT gestures to interlocutors in British English. Beattie and Shovelton also found that, whether in the presence of the original speech (Beattie & Shovelton, 2001) or isolated from it (Beattie & Shovelton, 2002), CVPT gestures are significantly “more communicative” than OVPT gestures on a number of measures concerning actions and/or objects: identity, description of action, shape, size, movement, direction, speech and relative position.

But this is only the tip of the proverbial iceberg: numerous studies have noted that changes in gestural viewpoint correlate with the emergence of deeper understanding, especially of abstract concepts — perhaps related to Beattie and Shovelton’s claim. In most circumstances, younger children and less advanced students seem to demonstrate a preference for CVPT gestures and only acquire OVPT gestures later (Gerofsky, 2010; Merola, 2009). These CVPT gestures let the child act out what would otherwise be an abstract situation. Because of this, some educators argue that the use of CVPT gestures drives the learning process and marks a first step in understanding (Petrick, Berland, & Martin, 2010).

In support of this, we may cite Wright (2001) who offers several examples of young students learning to reason abstractly about motion in a primary school classroom. She notes that one pair of students who found the task of observing and documenting changing motion particularly challenging were able to “get it” when they began enacting the changed motion events, using whole body enactments and CVPT gestures. They used their own movements across a “road” laid across their classroom to help solve their assigned problem sets. Others in the class who

had already mastered this basic inference had dropped the CVPT gestures and were instead using more schematic, trajectory-based OVPT gestures.

Similarly, Fadjo (2009) and Petrick et al. (2010) tracked a class of older students who were learning a basic programming language created for the purposes of their studies. The students' task was to individually program a soccer playing robot who could pair up with a random set of others in the class to compete as a team. Students who used problem-solving strategies which relied on CVPT gestures (such as embodying the robot and turning or kicking as it would) were shown to create more advanced, robust programs than other students who either did not gesture, or who used OVPT-based gesture strategies, such as rotating a finger to mark the robot's path. Even advanced researchers in theoretical physics were shown to reach better conclusions in abstract contexts when they were specifically asked to think through the reasoning process with a first person perspective which used CVPT gestures (Ochs, Jacoby, & Gonzales, 1994).

Together, these studies suggest that switching gesturing styles from one which emphasizes CVPT gestures (and by extension, reasoning from a first person perspective) to OVPT gestures indicates a shift in reasoning processes. Although the findings discussed here may seem contradictory, several studies have noted that changes in gestural viewpoint correlate with both a movement from understanding individual cases to their abstract, unifying structure (Gerofsky, 2010), as well as better memory of the taught concept (Merola, 2009). This relationship between changes in gestural viewpoint and depth of understanding is so strong that, as Gerofsky demonstrates, teachers are able to use it to predict student engagement and understanding with great accuracy.

These findings support the embodied cognition hypothesis (Lakoff & Johnson, 1980; Lakoff & Johnson, 1999), which claims that abstract meaning both develops from and depends on everyday reasoning with the body. They also support the claim by Parrill (2010b) that certain contexts prefer one viewpoint strategy over another. CVPT gestures seem to spark and develop understanding, especially in contexts which require abstract reasoning, while OVPT gestures seem to suggest mature understanding and the ability to work with abstract concepts schematically. When we consider this in the light of Fauconnier & Turner's (2002) claim that we understand abstract concepts by making them personal and human-sized, this is not so surprising a conclusion. We are used to reasoning with viewpointed bodies; adopting and extending those viewpoints, especially in circumstances which demand novel inferences and approaches, seems only appropriate.

Interim summary: Viewpoint strategies in co-speech gesture

The co-speech gesture strategies described here — facial expression and differences in gaze, CVPT vs. OVPT gestures, dual viewpoint gestures and other enactments — allow gesturers to embody characters and present narratives in vivid, dynamic ways. Although gestures usually express but one viewpoint, in many cases multiple viewpoints may also be expressed in gesture. Some factors, such as complexity of the described scene, the articulators involved or certain aspects of the accompanying speech, may affect the way viewpoint is expressed. In cases where multiple viewpoints are possible (such as the CVPT or OVPT construal of the robot's next soccer move), what is expressed in gesture may actually reflect depth of understanding and so have profound implications for education.

Viewpoint strategies in sign language

Unlike spoken languages where conceptual viewpoint is multimodal, largely inferred from linguistic input and supported by co-speech gesture (see Parrill, 2012, for a review), conceptual viewpoint in signed languages — whether inferred from linguistic or gestural information — takes place solely in the visuo-spatial modality. The expression of detailed information about the relative location, orientation, motion and activity of characters in an event is encoded from a physical perspective which the signs and their spatial organization embody (Özyürek & Perniss, 2011). Because of this visual flexibility, sign languages often encode detailed spatial information related to perspective (see Liddell, 2003, for a discussion of how it is included via gesture). In addition, because our bodies have relatively independent articulators, and because signers use space in a more structured way than do speakers (see Emmorey, 2002), signers are able to simultaneously express multiple physical and conceptual viewpoints in unique ways which speakers have yet to be shown to do. In this section, we will consider some of them. First we will consider the means by which viewpoint is expressed, and then we will consider situations in which preferences for one strategy or another have been noted.

Means of expression

In sign, as in co-speech gesture, each of the major articulators of the body may be used to express shifts in viewpoint. We describe each of the following in turn: gaze and facial information; manual signs (classifiers and handlers); simultaneous signs (constructions which simultaneously express multiple viewpoints across multiple

articulators); and constructed action/constructed dialogue, which uses the whole body.

Gaze and facial expression

Some signs require simultaneous articulation by the hands and face. In these cases, the face adds grammatical information. But there are other instances where the face is used gesturally to express conceptual viewpoint. The two examples described here illustrate how even subtle movements can profoundly affect conceptual viewpoint.

The first is the evidential marker in ASL (Shaffer, 2012), which is described via the “what I heard” blend. In this blend, typically, the signer uses the direction of gaze to suspend discourse and index the physical and mental space in which a previous discourse segment is reconstructed (Shaffer, 2012, pp. 152–153). The signer shows the reaction of their past self to the addressee: the signer turns her gaze to the side, signs “it was signed to me” there and shows her reaction to the reported utterance on her face. Subsequently, the signer actually signs her reaction and then, to end the account, returns her gaze to her actual addressee. In this way, the evidential marker is not only metonymic for witnessing the described event, but also reminiscent of face-to-face encounters, where participants are able to see each other react to events in real time.

The second example comes from Liddell (2003, pp. 208–209), who describes an utterance in ASL which would be translated into English as “[We will] compare the deaf and the hearing from the point of view of the Deaf.” The underlined portion of this utterance is not signed, but rather emerges from the signer’s use of physical space. In ASL, this utterance was produced as: DEAF HEARING. PEOPLE COMPARE-DUAL. Here, the signs for DEAF and HEARING were produced as the signer leaned her body to the right and then to the center-left, respectively, turning her face to match each lean. The sign for COMPARE was made from the right: the signer held her body in the real space for the Deaf and produced the sign COMPARE towards the real space for the hearing. Had she done the reverse, the English transcription would have read “...from the point of view of the hearing”, or had she maintained a neutral stance, no explicit point of view would have been mentioned. The perspective of the comparison emerged from the slight leans which the signer made in either direction and, more importantly, from the direction her face turned as each group was named. Other uses of the face and gaze to mark conceptual viewpoint interact with more complex strategies, and are described under the heading Constructed action/dialogue below.

Character vs. observer perspective

When describing co-speech gesture, character viewpoint (CVPT) and observer viewpoint (OVPT) refer to manual actions, specifically the degree to which the body is mapped in gesture space (partially or entirely for CVPT and not at all for OVPT). In the context of sign language, more precision is needed and so a distinction is made between the viewpoint adopted by the hands and the viewpoint adopted in sign space as a part of event projection.

Manual perspective largely rests on handshape distinctions, the most common division of which involves forms which depict objects and forms which depict the handling of objects, commonly called ENTITIES and HANDLERS (see Schembri, 2003, for a review). ENTITIES typically represent whole or part of the referent, and are based on salient physical characteristics of size and shape which create iconic shape representations, e.g., the /V/ handshape in British Sign Language which can be used to represent scissors (Cormier, Quinto-Pozos, Sevcikova, & Schembri, 2012). In this sign, the index and middle fingers of the dominant hand are displayed in a “V” shape, like the emblem for “peace” or “victory”. HANDLERS, on the other hand, typically represent the way in which whole or part of the referent is handled or touched, e.g., the /5/ handshape in British Sign Language (displayed with wiggling fingers) used to denote playing the piano (Cormier et al., 2012). In this sign, the palm of the dominant hand is held up and all five fingers are spread, wiggling, showing how one might play the keys on a piano. This two-way distinction nicely parallels the CVPT vs. OVPT distinction made in gesture (McNeill, 1992).

In sign languages, we also find character perspective (CVPT) or observer perspective (OVPT). Although other terms are possible (see Özyürek & Perniss, 2011 for a review), we follow Perniss’ (2007a) nomenclature since she makes an explicit, purposeful link between sign and gesture terminology. According to her, CVPT space denotes an event space which is projected onto sign space from a character’s physical viewpoint within an event while OVPT space denotes an event space which is projected onto sign space from an external physical viewpoint in front of the signer, e.g. from the kind of third person perspective offered by most narrators (Perniss, 2007a). Prototypical alignments occur when both manual and spatial perspectives are the same; non-alignments occur when they differ, such as when a HANDLER (a CVPT sign) occurs in OVPT space. Non-alignments will be discussed in the subsection on simultaneous signs and body partitions.

Aligned event spaces come in two types: CVPT space with HANDLERS (CVPT signs) and OVPT space with ENTITIES (OVPT signs). The first is a case of true enactment since the signer’s own body movements and local space map onto those of the character. An example of this, from German Sign Language, is a signer retelling a cartoon in which a mouse plays with a ball: the signer shows

an imagined life-sized ball in front of her body and moves her arms to demonstrate throwing it as the mouse did, forward along the sagittal axis (Perniss, 2007b, p. 1331). In the case of aligned OVPT space, entities are represented from afar, as they were witnessed. In the same cartoon as before, the mouse stands across from an elephant; they are shown to the left and right of each other on screen. In German Sign Language, the signer represents this with an ENTITY classifier for each animal on either side of her body (Perniss, 2007a: p. 203). The signer doesn't map her head or torso into the event space — only her hands represent entities in the reported situation.

These prototypical alignments nicely parallel the discussion of CVPT and OVPT gestures in the previous section, and point to the fact that we seem to be able to take two basic embodied perspectives on events, as a participant or observer of an event.

Constructed action

Constructed action is the selective reconstruction of a referent's actions or thoughts (Liddell & Metzger, 1998; Metzger, 1995). It has also been called "constructed dialogue" (Quinto-Pozos, 2007), "referential shift" (Emmorey & Reilly, 1995), "quotative role shift" (Pfau & Quer, 2010) and "role shift" (Lillo-Martin, 1995; Padden, 1990). It includes representations of actions, utterances, thoughts, attitudes and feelings, etc. and is a widespread feature of diverse sign languages. It is the means by which signers can embody other animate entities, including themselves in another time or place.

Although there is some evidence that constructed action should be sub-divided into different categories (e.g., Emmorey & Reilly, 1998, who show that constructed dialogue is used by age 7 but constructed action takes longer), most recent discussions focus on constructed action as a single category which allows for mimetic representation akin to the demonstrations discussed by Clark & Gerrig (1990). Perhaps because of this, constructed action is one of the most productive constructions in sign languages, and is widely reported in ASL and its relatives, as well as in Israeli Sign Language, Irish Sign Language and Taiwanese Sign Language, to name a few (see Janzen, 2004, for a review of constructed action, and Quinto-Pozos, 2007, for a discussion of its status as a grammaticalized feature of signed languages).

Constructed action is marked by the systematic use of the signer's head, face, hands and body to enact a given character, thereby allowing the signer to show the adoption of another's viewpoint. It ranges from subtle enactments, which use the face and gaze, to elaborate enactments, which also use the torso to depict CVPT events in CVPT space (Perniss, 2007a; Quinto-Pozos, 2007). It usually begins with a sign that identifies the referent (e.g., by full NP or fingerspelling the name),

followed by disengaging eye contact with the signer's actual addressees and shifting gaze and body posture off-center — a shift typically made in horizontal space. Once oriented to a new space, the signer can then sign from that character's perspective, just as speakers of oral languages can quote one another's speech. Each direction, and thus, each change in physical viewpoint, corresponds to another character in the signer's narrative. For example, a signer who quotes a small child in ASL would direct her own signs upward, as if she were a small child signing to an adult (Taub, 2001). If that narrative were to continue and the signer next quoted an adult signing to the child, the signs would be directed downwards, to an invisible child-sized addressee.

Sign language researchers typically point to constructed action as “the way” in which character viewpoint is expressed in sign language discourse, but as Janzen (2004, 2012) points out, it is actually one option which seems to be preferentially used to highlight comparison. Another option, described by Janzen (2012) as rotating the vantage point, creates rotated spaces and is analyzed in terms of a network of blends. As with constructed action, the signer signs from the quoted/enacted character's perspective. However, unlike constructed action, the spatial distribution of characters is virtual rather than real. Because of this, better terms for these strategies may be body-rotation role shift or virtual-rotation role shift.²

When rotating the vantage point, a signer embodies first one character and then another without body rotation. Instead, the real space blending shifts so that the space in front of the signer becomes, in turn, the space in front of each character. Frequently the two characters are facing each other so there is a 180 degree rotation of viewpoint of the imagined situation. Only the real space and the spatial displacement of referents suggest this difference. There are no other physical cues that a viewpoint shift has taken place. Thus, rotating the vantage point requires that the signer and the addressee actively maintain each participant's conceptual viewpoint in memory, and rotate the real space to match it for each character change. Moreover, unlike constructed action, the circular rotation inherent to rotating the vantage point requires that each character establish the relevant location of other characters from their physical viewpoint each time.

Typically, signers rotate space by 180 degrees, since most conversations, especially signed ones, happen face-to-face. As Janzen points out, space is rotated and not simply reversed. For example, two characters active within a rotating the vantage point network may point to the same physical location (e.g., to the right of the signer's body) and intend different referents (such as JUDY and HER.DOG) since each character maintains his own physical viewpoint which is aligned with the signer's.

Simultaneous signs/body partitions

The simultaneous expression of multiple viewpoints is a frequent strategy which maximizes information content, allowing addressees to experience a given scene from multiple physical viewpoints at once. Sometimes, as in the case of the ASL evidential marker described earlier (Shaffer, 2012), the face and head express the signer's reaction to the content of the signing space elaborated by the hands. But strategies which partition the body into multiple physical viewpoints also exist. We will consider each of these in turn.

Dudis (2004) describes several possible partitions of the body for the simultaneous expression of physical viewpoint. Similar to Parrill's (2009) description of dual viewpoint gestures, the articulators are seamlessly divided amongst different characters in an event. For example, Janzen describes an example from ASL in which the signer's left arm and face represent a cop while the right arm represents the signer herself during that encounter (2004). Aarons and Morgan (2003) describe several instances of this kind of partitioning in South African Sign Language, where multiple simultaneous viewpoints commonly co-occur with classifiers. They claim that the distribution of viewpoint across articulators reinforces both perspective and topic maintenance, and thus is a creative and efficient meaning making strategy which is capable not only of tracking multiple perspectives, but of doing so simultaneously. One example they give (p. 131) is of a signer describing a picture in which a parachutist drifts past an airplane: Having already shown the parachutist drifting through space with a series of CVPT signs, the signer partitions his body such that the right hand signs the classifier for airplane and maintains the parachutist's location in relation to it while the signer's face and left hand sign the parachutist's first person reaction to it. In terms of Perniss (2007b), this constitutes a non-aligned use of event space: The event space is CVPT while the signer's body is partitioned and gives both CVPT and OVPT information.

Partitioning the body in this way, in order to show the simultaneous interaction of multiple entities, is a common strategy in sign languages. Showing the simultaneous interaction of multiple spaces is much less common and has, to date, only been documented in Turkish Sign Language in what Perniss & Özyürek (2008) call fused perspective. In fused perspective, the signer simultaneously expresses both CVPT and OVPT perspectives: they map the character's head and torso onto their body, moving those articulators from the character's perspective, but project the movements into signing space from an observer's perspective. As an example, participants in their study watched a film in which a mouse makes a pancake, flips it — and misses, accidentally dropping the pancake on the floor. On the screen, this movement is horizontal but within the narrative space, the movement is across the sagittal axis of the characters (that is, made in front of the

mouse's body). When signing this story to others, Özyürek & Perniss's participants signed from the perspective of the mouse, showing the frying pan in front of the body, but flipped the pancake on the lateral axis. This manual action matched the way the signers had witnessed the event in the cartoon, not the way the mouse within the story experienced it, even as the gaze, head and shoulders of the signers continued to describe the mouse's actions.

Descriptions like this conflate observer and character viewpoints, and somehow represent better what the signers saw on screen than either a purely OVPT or purely CVPT retelling would. Because of this, Perniss & Özyürek argue that the construction increases efficiency and informativity since it combines the signer's and the character's respective conceptual viewpoints.

Less dramatic examples of this combination are Perniss' non-prototypical alignments, which have been documented in German Sign Language and Turkish Sign Language (2007b). These are cases in which the signer projects multiple perspectives onto the signing space: either CVPT space with OVPT classifiers or OVPT space with CVPT classifiers. The first case is exemplified by the South African Sign Language data mentioned above. The second case is exemplified by an example from Perniss & Özyürek (2008, example 2, p. 363): the signer's hands simultaneously represent participants (CVPT classifiers) and the spatial relations between them (OVPT space) in the described event. This combination allows for discourse coherence, e.g., concerning the efficiency of perspective sharing which is supported by the finding that simultaneous constructions reduced the number of utterances needed to describe a cartoon by half (Perniss, 2007b), or the ability to express unambiguous information about precise spatial relationships.

This expression of multiple viewpoints is visually complex, and may reflect both properties of the original stimulus and the signer's visual experience of them. In both cases, it increases the information content of the utterance by condensing multiple event perspectives into a single utterance.

Viewpoint preferences

As we have seen, signers have a number of options for expressing differences in conceptual viewpoint, most of which hinge on the difference between CVPT and OVPT. But what are the circumstances under which one or the other kind of expression tends to occur?

Linguistic preferences

Perniss (2007) and Özyürek and Perniss (2011) show that in sign languages, viewpointed spatial information is conveyed in two ways (manual signs and event space projection) which may combine to be prototypically aligned or

non-aligned. Moreover, they suggest that these viewpoint combinations reflect affordances of the visuo-spatial modality. In their comparison of German Sign Language to Turkish Sign Language, Özyürek and Perniss find a strong preference for viewpointed predicates over lexical predicates, with a strong preference for HANDLERS (CVPT signs) over ENTITIES (OVPT signs), and a preference for CVPT space over OVPT space. This points to an overall preference to embody characters and act from their perspective which is in line with observations by Engberg-Pedersen (1993) on Danish Sign Language, Earis & Cormier (2010) on British Sign Language, and Quinto-Pozos (2007) and Rayman (1999) on American Sign Language.

The degree to which each language uses CVPT can of course differ (e.g., Özyürek & Perniss, 2011, notes that German Sign Language uses CVPT space more often than Turkish Sign Language does, and suggests that this might be because German Sign Language has an explicit CA marker), but the overall tendencies remain. Along the same lines, aligned perspectives, in which both the event space projection and the manual signs reflect the same viewpoint, are more common than unaligned perspectives, with individual languages determining the relative frequencies of each. Turkish Sign Language seems particularly adept at combining information from multiple physical viewpoints (Perniss & Özyürek, 2008).

As with Parrill (2010b) in her work on co-speech gesture, Özyürek and Perniss (2011), Zwitserlood (2003) and Engberg-Pedersen (1993) found that transitive utterances in signed languages tend to be accompanied by CVPT signs while intransitive utterances tend to be accompanied by OVPT signs. Earlier, we mentioned that this distinction is motivated by iconic properties, such as which body parts are involved, as actions or objects which involve the hands seem particularly suited to CVPT perspective. In addition, Özyürek and Perniss (2011) found that event type can predict the use of aligned vs. non-aligned perspectives: events involving one animate entity are more likely to be depicted with aligned perspectives whereas events involving multiple animate entities, or an animate entity with an object, are more likely to be depicted with non-aligned perspectives, thus showing that physical viewpoint, linguistic viewpoint and the semantics of the event interact to determine representation.

Task-specific preferences

In general, when asked to retell narratives, signers show a strong preference for telling them from the perspective of the characters involved using constructed action — regardless of whether the original narrative was presented with a written description (Earis, 2008) or a short cartoon (Rayman, 1999). Considering how prevalent constructed action is, Quinto-Pozos (2007) posed the question: is it obligatory?

To answer that question, Quinto-Pozos asked 10 native/early signers of American Sign Language to watch 20 short clips (of animate and inanimate subjects) and retell them twice. The first time, participants were asked to describe the clips spontaneously, using whatever means seemed the most natural to them. In the second retelling, the experimenter asked them to retell the same clip, but with one crucial difference, e.g., if they used constructed action the first time, they were asked to not use it the second time. Quinto-Pozos found that when participants were asked to omit constructed action, nearly 50% of them refused to do so, claiming it was impossible. A few participants opted instead to increase the complexity of the conceptual viewpoint. For example, instead of representing a bee landing on a flower directly, they elaborated the original clip and showed somebody watching the bee buzz around before landing on the flower. When asked about their choice for describing the clips, most participants reported that constructed action let them do so in a more clear and precise way.

In a follow-up study, Quinto-Pozos asked 18 native/early signers of American Sign Language to watch the same clips and the retellings the first group of signers had produced. They were asked to rate each retelling for clarity and correctness. Quinto-Pozos found that the two tellings were rated very differently. In general, and following the instincts of signers in the first study, he found that the use of constructed action was preferred, with instances with constructed action rated as being both more clear and more correct. In practice, this meant that the first spontaneous telling was preferred — a few cases resulted in a preference for the second telling, but in those cases, the second telling introduced constructed action. This finding lends strong support to Özyürek and Perniss's (2011) observation that CVPT perspective is preferred.

In a related study, McCleary and Viotti (2009) conducted a qualitative micro-analysis of a narrative in which they found that one signer reliably maintained the conceptual and gestural viewpoint of the original stimulus when retelling it later (there are hints of other signers maintaining viewpoint in Earis, 2008, and Rayman, 1999). Working with native signers of Brazilian Sign Language, McCleary and Viotti describe one narrative in particular in which the signer maintained the perspective of the original stimulus, faithfully reproducing the conceptual viewpoint from which it was viewed. In their study, McCleary and Viotti showed a signer the Pear Story (Chafe, 1980) and asked him to retell it. When he did, they discovered that the viewpoint of the shots used in the film corresponded exactly to those used in the signer's narrative: scene-by-scene, the depicted viewpoints and actions matched. This lends support to the idea that signers are not only sensitive to viewpoint, but that they go so far as to reflect the continuity and identifiable shifts in point of view that they experience.³

Interim summary: Viewpoint strategies in signed language

Expanding McNeill's (1992) terminology to signed languages as Perniss and Özyürek (2008) and Cormier et al. (2012) do makes it easier to see how affordances of the visuo-spatial modality are similarly reflected in linguistic communities, regardless of the primary mode of communication. Although there may be more terminological discussions in the sign literature, the basic CVPT/OVPT division remains, with signers showing a strong preference for CVPT depictions. Signs display viewpoint choices with regards to manual signs (HANDLERS and ENTITES, roughly CVPT and OVPT signs) and spatial projection (CVPT and OVPT), which interact to create the aligned and non-aligned event projections described by Perniss and colleagues. This choice seems to be motivated by a need for clarity, precision and correctness (Quinto-Pozos, 2007), but is sensitive to features such as animacy, transitivity, and number of entities involved (Özyürek & Perniss, 2011). Signed utterances which conflate multiple viewpoints, such as non-aligned event projections (Perniss, 2007), fused perspective (Perniss & Özyürek, 2008), and body partitions (Dudis, 2004) seem to be more widespread than in co-speech gesture, and perhaps reflect the flexibility signers have with the visuo-spatial modality.

A comparison of viewpoint strategies in gesture and sign language

Because the best representation of a viewpointed body is another viewpointed body (Sweetser, 2012), it should come as no surprise that both co-speech gesture and sign language use the body in similar ways to iconically represent conceptual viewpoint. There are of course important differences between these two groups, but as we have already seen, bodily representations of viewpoint for each come in two basic forms: CVPT and OVPT. To better understand how they are used, and what the differences in their use are, we will now compare viewpoint strategies in co-speech gesture and sign. We begin with manual behaviors. Then, we look to viewpoint preferences in narrative and spatial reasoning tasks. Finally, we present a table which documents the viewpoint strategies used in both groups. Taken together, these show that although most viewpoint strategies employing visible bodily action are available and widely documented in both groups, the way in which each group uses them is very different.

Two strategies: CVPT vs. OVPT

Earlier, we saw that gesturers can use their hands to show objects with CVPT or OVPT (McNeill, 1992), and that signers also make this distinction, describing

objects with either HANDLERS or ENTITY classifiers (Schembri, 2003). Cormier et al. (2012) make explicit connections between CVPT gestures and HANDLERS and OVPT gestures and ENTITIES, claiming that there are quantitative differences but not qualitative ones; the production of each class is similar in both groups. In particular, they note the finding of Schembri, Jones, & Burnham (2005) who compared native speakers of Australian English to native/early signers of Australian Sign Language. Schembri et al. found that on the verbs of motion task, speakers and signers express similar OVPT information in similar ways with regards to movement and spatial arrangement of the hands, and differ only with regard to handshape, which is a grammatical feature of sign.

Schembri et al.'s finding is in line with the findings of Brentari, Coppola, Mazzoni, & Goldin-Meadow (2011). In their study, Brentari et al. compared two groups of signers (Italian Sign Language and American Sign Language) against two gesturing-only groups composed of speaker-hearers (Italian and English) in two age groups (children aged 4–8 and adults) and with both gestural viewpoints (CVPT/HANDLERS and OVPT/ENTITIES). The speaker-hearers were asked to pantomime — i.e., to gesture without speech. Participants were shown clips of objects and people handling those objects, and were asked to reproduce them for a partner. They found that, overall, the signers patterned similarly and the speakers patterned similarly, regardless of language or age. With respect to viewpoint, they found that CVPT/HANDLERS were expressed in similar ways, which may be due to the fact that the speakers simply imitated the object manipulation they had seen, and this handling nicely matched the CA used by the signers. There was much more variation with the OVPT/ENTITIES group, mostly based on handshape complexity. Brentari et al. hypothesize that this is because the signers were influenced by the conventionalized handshape categories which their languages provide — unsurprising since the typical speaker-hearer would probably not know the classifiers of the local signed language.

Like Schembri et al. (2005), Brentari et al. (2011) found a number of similarities between signers and speakers, as well as differences which stem from grammaticalized or conventionalized forms. Based on this and the narrative preferences discussed below, Cormier et al. (2012) propose that the CVPT/OVPT distinction which pervades gesture studies should also be used in sign studies in order to more easily draw parallels between uses of the visuo-spatial modality. Perniss (2007) and Sweetser (2012) make a similar argument.

Viewpoint preferences

Both speakers and signers exhibit preferences for particular viewpoint strategies in different contexts. In this subsection we consider two, narratives and spatial reasoning.

Viewpoint preferences in narrative

Earis and Cormier (2010), Marentette, Tuck, Nicoladis, & Pika (2004), and Rayman (1999) document differences between speakers and signers with regard to viewpoint preference and narrative strategy. These differences appear to stem from the fact that storytelling could evidence a cultural difference (Peters, 2000): Signers tend to be skilled storytellers who focus on detailed information from a character's perspective whereas speakers tend to prefer narrator perspectives which suggest narrative detail. This holds for both quantitative and qualitative measures. For example, Marentette et al. (2004) asked participants to retell a fable, and compared ASL signers (native signers vs. late signers vs. English/ASL bilinguals) against monolingual speakers of English. They found that the average length of stories was longer in ASL than in spoken English, and that constructed action was used much more often than CVPT gestures. Even amongst the signers, these observations held, with native/early signers taking more time and using more constructed action to retell the fable than late signers or English bilinguals did.

In terms of qualitative analysis, Rayman (1999) asked 5 native speakers of English and 5 native signers of ASL to retell the same narrative (the tortoise and the hare fable) from a cartoon stimulus. Participants were matched on all demographic measures, including exposure to theatre. Rayman found that English speakers tended to tell short stories from the narrator's perspective, and relied on their addressee's ability to infer story details such as the fact that the two animals began their race with a gunshot aimed to the sky. Most of her English-speaking participants did not gesture (Rayman, 1999, p. 72). One participant, trained in theatre, did gesture and was quite animate in comparison to the others, but her narrative was still shorter than the average signed story, and she used visible bodily action to characterize the main characters to a lesser extent than the signers. In the same task, Rayman found that signers tended to use constructed action to give rich spatial and visual information. Their stories were longer and were told from each character's perspective, focusing on details and supporting actions to a greater extent than did the speakers.

In a follow-up study, using the same fable, Earis and Cormier (2010) conducted a micro-analysis to compare the viewpoint preferences of native British Sign Language signers against two expert storytellers ("speakers") from the Scottish Storytelling Center. Earis and Cormier also found that speakers tell shorter stories

from a narrator's perspective compared to signers, who give more detailed stories from each character's perspective. The speakers in this study, however, did tend to use CVPT gestures in creative ways, often supplementing them with vocal gestures which gave an overall effect similar to constructed action (though without its structural properties, such as the consistent use of different spaces for different characters). As expected, the signers used constructed action to structure their narratives, assigning each character to a consistently-used space. In addition, Earis and Cormier also found that speakers tended to point to characters in the same way that signers index characters. Although multiple studies agree on the behavior of signers, we have too little information on the behavior of speakers in comparable tasks. Taken together, these studies seem to confirm an earlier finding that input modality does not affect gestural viewpoint produced (Parrill, Bullen, & Hoburg, 2010). They do, however, suggest that viewpoint choices may be an effect of production modality — but to what extent remains to be seen.

Viewpoint preference in spatial tasks

If we turn now to spatial reasoning tasks, we also find a difference in viewpoint preference. Spatial reasoning tasks ask participants to navigate through space, e.g., by giving a route from one landmark to the next using a real world location (such as San Francisco) or an invented one, such as the small town used in Park-Doob (2010). Typically, participants use one of two options: a survey perspective (OVPT) which takes a bird's eye view of the scene; or a route perspective (CVPT) which structures navigation from a particular first person point of view. Considering these two options, and the fact that signers have demonstrated such a strong preference for CVPT in other tasks, the question naturally arises: how do signers and speakers compare when navigating through space?

Emmorey, Tversky, and Taylor (2000) answered that question by asking groups of native or early signers⁴ of American Sign Language and native speakers of English to complete a spatial reasoning task. They found that each group presented very strong, very different preferences for solving it. Signers demonstrated a preference for depicting highly schematic scenes which used OVPT: descriptions traced the trajectory of a person's movement across a map of the area which projected into the space in front of the signer. In contrast, English speakers preferred projections from "within" the scene (roughly, CVPT gestures), which involved left and right points and torso shifts, and no projection of a map into the space in front of them.

This is a curious reversal of the narrative pattern, which saw signers demonstrate a strong preference for CVPT construals. Those construals were made in order to maximize information content, especially clarity, since all relevant aspects of the scene could be presented simultaneously (Quinto-Pozos, 2007). The

same can be argued here: a 2D projection of a map into the signer’s lap (or, into the shared speaker-addressee space) where visible landmarks and routes can be explicitly marked is much more clear than a series of left or right turns.

Interim summary: Viewpoint strategies in co-speech gesture and signed languages

Considering these findings, and those described in previous sections, we can now summarize the strategies using visible bodily actions that are used to encode differences in physical and conceptual viewpoint (Table 1).

Table 1. Summary of bi-modal viewpoint markers in co-speech gesture and sign language

Domain	Marker	Description	Features
Gesture	Character VPT gestures	Sub-class of iconic gestures. Preferentially used to indicate emotional states, manner, role shift, and activities performed with the torso or hands from the enacted character’s point of view. Often used accompanying transitive utterances (McNeill, 1992; Parrill, 2010).	May be a preferred language-specific gesture strategy (e.g., Brown, 2008) which conveys more information content to addressees (Beattie & Shovelton, 2002). Conflicting accounts as to whether it indicates a deeper, mature understanding of conceptual information (Ochs et al., 1994; Petrick et al., 2010) or a more basic understanding (Wright, 2001). Some indication that CVPT is used first when confronted with abstract concepts, e.g., Merola (2009).
	Observer VPT gestures	Sub-class of iconic gestures. Preferentially used to indicate path, manner+path and spatial layout, or when participants have shared knowledge or common ground. Typically gives third person perspectives of scenes or schematic depictions of events, like trajectory for motion events. Often used accompanying intransitive utterances (McNeill, 1992; Parrill, 2010).	May be a preferred language-specific gesture strategy (e.g., Brown, 2008). Conflicting accounts as to whether it indicates a deeper, mature understanding of conceptual information (Wright, 2001) or a more basic understanding (Ochs et al., 1994; Petrick et al., 2010). Some indication that OVPT is used when abstract concepts are understood, e.g., Merola (2009).
	Gesture Location	Indicates the spatial location of entities, especially in real space blends, which pair spatial locations with conceptual information. Often indirectly indicates conceptual viewpoint.	In combination with body orientation, face orientation, gesture location or eye gaze, gesture location may indicate role shift , character viewpoint or a rotated space (e.g., Earis, 2008).
Sign Language	Handlers (CVPT signs)	Type of classifier. Typically used to indicate the way in which the referent is handled or touched (Schembri, 2003). May indicate whole or part of the referent. Often occur in transitive utterances, especially with animate entities (Özyürek & Perniss, 2011).	“Handlers” have many names — see Schembri (2003) for a review. Display similar iconicity, and sensitivity to linguistic constructions, as CVPT gestures (Cormier et al., under review).

Table 1. (continued)

Domain	Marker	Description	Features
	Entities (OVPT signs)	Type of classifier. Typically used to indicate schematic features of the referent which are based on salient physical characteristics of size and shape (Schembri, 2003). May indicate whole or part of the referent. Often occur in intransitive utterances (Özyürek & Perniss, 2011).	“Entities” have many names — see Schembri (2003) for a review. Display similar iconicity, and sensitivity to linguistic constructions, as OVPT gestures (Cormier et al., under review).
	Aligned event space projections	An event space projection in which the projected space and the manual signs depict the same viewpoint: either CVPT or OVPT (Perniss, 2007b). Typically occur when single entities are involved, e.g., with intransitive utterances (Özyürek & Perniss, 2011). More frequent than non-aligned projections (Perniss & Özyürek, 2008).	Aligned CVPT spaces are instances of true embodiment, i.e., Constructed Action/Role Shift (Perniss, 2007; Quinto-Pozos, 2007).
	Non-aligned event space projections	An event space projection in which the projected space and the manual signs depict different viewpoints, e.g., Handlers in OVPT space, or Entities in CVPT space (Perniss, 2007b). Typically occur when multiple entities and/or objects are involved, e.g., with transitive utterances (Özyürek & Perniss, 2011).	Non-aligned perspectives maximize information content by compressing multiple physical viewpoints into one space (Özyürek & Perniss, 2011; Perniss & Özyürek, 2008).
	Rotating Spaces	Indicates a different conceptual viewpoint which is established by gesture location and orientation. May also be established by eye gaze or face and body orientation (Janzen, 2012; Janzen, 2004).	The real space is divided amongst conceptually present characters in the narrative, whose relative positions are maintained as the signer shifts between them (Janzen, 2012; Janzen, 2004). Often accompanied by character-viewpointed gestures and role shifted signing (Janzen, 2012).
Both	Body Partitioning	Typically includes a division of the body, such that two or more distinct aspects of an event are simultaneously represented in different ways on different parts of the same body. The torso may be divided from the head, the left side of the body from the right, etc. (e.g., Dudis, 2004; Parrill, 2009). Typically includes the simultaneous expression of gestures or signs from multiple physical (Perniss & Özyürek, 2008; Narayan, 2012) or gestural (Parrill, 2009) viewpoints.	Indicates most commonly the simultaneous expression of multiple viewpoints (dual viewpoint gestures or simultaneous signs), e.g., Aarons & Morgan (2003); Özyürek & Perniss (2011); Parrill (2009); Perniss & Özyürek (2008).

Table 1. (*continued*)

Domain	Marker	Description	Features
	Constructed Action (Role Shift)	Indicates a different conceptual viewpoint which belongs to another character or to the speaker in another place and time (e.g., Earis, 2008; Quinto-Pozos, 2007). Typically indicated by non-addressee-oriented face and body orientation and eye gaze (e.g., Quinto-Pozos, 2007). Often, the speaker's face takes on characteristic features of the depicted character, such as affective displays (e.g., Earis & Cormier, 2010; Rayman, 1999).	Often accompanied by character-viewpoint gestures or signs (e.g., Earis, 2008).
	Eye Gaze	Indicates the transition from one conceptual viewpoint to another, or from one mental space to another (e.g., Maury-Rouan, 2011). Typically includes disengaging (or re-engaging) eye gaze from the addressee and redirecting it to another space (e.g., Shaffer, 2012). Typically accompanied by changes in body and face orientation, or differences in pitch-accent (e.g., Earis & Cormier, 2010)	When combined with non-addressee-oriented face and body orientation, eye gaze may indicate role shift or character viewpoint . When combined with addressee-oriented face and body orientation, it may indicate a rotated space or discourse-level viewpoint (e.g., Bavelas & Chovil, 2000; Janzen, 2012; Maury-Rouan, 2011; Shaffer, 2012).
	Body Orientation	Used for physically representing and locating different conceptual viewpoints, e.g., Janzen (2012) or Liddell (2003). Typically includes shifting or rotating the torso and directing gestures to a new signing space. This space/body orientation then becomes a real space blend for the conceptual viewpoint located there (e.g., Janzen, 2012).	When combined with eye gaze, body orientation may indicate role shift or character viewpoint (e.g., Earis & Cormier, 2010; Fadjo et al., 2010; Janzen, 2012).
	Facial Information	Preferentially used for identifying different conceptual viewpoints and displaying character viewpoint (Fox & Robles, 2010; McClave, 2000).	When combined with eye gaze, face orientation or expression may indicate role shift or character viewpoint (e.g., Bavelas & Chovil, 2000; Earis & Cormier, 2010; McClave, 2000; Shaffer, 2012).

Summary and suggestions for further inquiry

The focus of this review has been on the strategies used in co-speech gesture and sign languages to express differences in conceptual viewpoint, especially with

regards to the strategies and affordances offered by using gesture. Among the strategies found to convey differences in viewpoint were: facial displays, direction of gaze, orientation and gesture. These were considered in terms of both descriptive and experimental work, the combination of which suggests that although the strategies used in sign language are more structured than those found in co-speech gesture, the *kinds* of strategies used and the ways in which they combine with linguistic means to express conceptual viewpoint are similar. In particular, the division of iconic gestures into character vs. observer viewpoint gestures (handling vs. entity constructions in sign), and the situations in which one or the other tends to appear, is remarkably similar for both co-speech gesture and sign languages. Likewise, the ability to partition the body to simultaneously represent multiple viewpoints was also shown to be in common use in both communities.

Of course, a number of questions still remain, and there is plenty of room for further research. For example, the relationship between shared knowledge or common ground and gestural viewpoint needs to be explored. So too does the degree to which real space blends and local spatial organization affect how viewpoint is expressed in gesture. Finally, as Bavelas and colleagues (Bavelas, Beavin, & Chovil, 2000; Bavelas & Gerwing, 2007) point out, a systematic study of the communicative uses of facial expressions (affective or otherwise) and gaze needs to be carried out. Thanks to the work of Shaffer (2012) and others, we are beginning to understand their significance in the grammar of different sign languages, but comparative work done in speech communities is lacking.

Although the summary of documented strategies using visible bodily action used to convey viewpoint (Table 1) might easily look as if it claims to be an absolute list of such viewpoint expression, people often prove to be adept at creating and constructing meaning in even the most meagre of circumstances. We are very sensitive to even the most subtle visible action or shifts in conceptual viewpoint; in theory, any gesture or change in gesturing style may potentially mark a viewpoint shift, just as any linguistic cue or shift in accent may serve to mark one. The literature reviewed here demonstrates that we are constantly inhabited by other bodies, including their affordances and physical viewpoints. We cannot help but track and communicate them. We are used to tracking the points of view of those who share our physical spaces. It only makes sense that this skill, like any other general cognitive strategy, be part of an everyday communicative process.

Acknowledgements

We thank Mike Huiskes, Alan Cienki, Gisela Redeker, Eve Sweetser and members of the Berkeley Gesture Group, an anonymous reviewer and the Editor of *Gesture* for helpful comments made

on earlier versions of this paper. This work is part of the research program *The Conversation Frame: Linguistic Forms and Communicative Functions in Discourse*, which was awarded to Dr. Esther Pascual by the Netherlands Organization for Scientific Research (NWO).

Notes

1. Of course, speakers use gestures in a number of ways. Although we would argue that all gestures are necessarily viewpointed, it is of course the case that some gestures, such as iconics, are more easily described as being viewpointed than others (e.g., points, beats, or pragmatic gestures).
2. Thanks to Eve Sweetser for these terms.
3. We do not yet know whether speakers are also able to reproduce viewpoint shifts to the same degree following presentation of a video stimulus. But see Kendon (2004, ch. 9) for a detailed account of an example in which one sees fluent shifts in viewpoint.
4. In this study, native signers [$n=27$] were raised from birth in an ASL environment; early signers [$n=10$] learned ASL before age 7 but missed the early exposure to sign the native signers had, and 3 participants learned ASL after age 14.

References

- Aarons, Debra & Ruth Zilla Morgan (2003). Classifier predicates and the creation of multiple perspectives in South African Sign Language. *Sign Language Studies*, 3 (2), 125–156.
- Alibali, Martha W. & Sotaro Kita (2010). Gesture highlights perceptually present information for speakers. *Gesture*, 10 (1), 3–28.
- Alibali, Martha W., Robert C. Spencer, Lucy Knox, & Sotaro Kita (2011). Spontaneous gestures influence strategy choices in problem solving. *Psychological Science*, 22 (9), 1138–1144.
- Barsalou, Lawrence W. (1999). Perceptual symbol systems. *Behavioral and Brain Sciences*, 22, 577–609.
- Bavelas, Janet Beavin & Nicole Chovil (2000). Visible acts of meaning: An integrated message model of language in face-to-face dialogue. *Journal of Language and Social Psychology*, 19 (2), 163–194.
- Bavelas, Janet & Jennifer Gerwing (2007). Conversational hand gestures and facial displays in face-to-face dialogue. In Klaus Fielder (Ed.), *Social communication* (pp.283–307). New York: Psychology Press.
- Beattie, George & Heather Shovelton (2001). An experimental investigation of the role of different types of iconic gesture in communication: A semantic feature approach. *Gesture*, 1 (2), 129–149.
- 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.

- Bergen, Benjamin K. (2012). *Louder than words: The new science of how the mind makes meaning*. New York: Basic Books.
- Brentari, Diane, Marie Coppola, Laura Mazzoni, & Susan Goldin-Meadow (2011). When does a system become phonological? Handshape production in gesturers, signers, and homesigners. *Natural Language and Linguistic Theory*, 30, 1–31.
- Brown, Amanda (2008). Gesture viewpoint in Japanese and English. *Gesture*, 8 (2), 256–276.
- Bundgaard, Peer (2010). Means of meaning making in literary art: Focalization, mode of narration and granularity. *Acta Linguistica Hafniensia*, 42, 64–84.
- Chafe, Wallace L. (1980). *The pear stories: Cognitive, cultural, and linguistic aspects of narrative production*. New York: Ablex.
- Chafe, Wallace L. (1994). *Discourse, consciousness and time*. Chicago: University of Chicago Press.
- Clark, Herbert H. & Richard J. Gerig (1990). Quotations as demonstrations. *Language*, 66 (4), 764–805.
- Cormier, Kearsy, David Quinto-Pozos, Zed Sevcikova, & Adam Schembri (2012). Lexicalisation and de-lexicalisation processes in sign languages: Comparing depicting constructions and viewpoint gestures. *Language and Communication*, 32 (4), 329–348.
- Dudis, Paul (2004). Body partitioning and real-space blends. *Cognitive Linguistics*, 15 (2), 223–238.
- Dudis, Paul (2007). *Types of depiction in ASL*. Unpublished manuscript. Retrieved from: <http://www.gallaudet.edu/documents/academic/drl-dudis2007.pdf>, August 14, 2012.
- Earis, Helen (2008). Point of view in narrative discourse: A comparison of British sign language and spoken English. Unpublished doctoral dissertation, University College London. Retrieved from: <http://kearsy.macmate.me/web/Earis.html>, August 14, 2012.
- Earis, Helen & Kearsy Cormier (2010). Point of view in British Sign Language and spoken English narrative discourse: The example of 'The Tortoise and the Hare'. Unpublished manuscript. Retrieved from: http://kearsy.macmate.me/web/Publications_files/Earis_Cormier_submitted7Jan2010.pdf, August 14, 2012.
- Emmorey, Karen (2002). *Language, cognition, and the brain: Insights from sign language research*. Mahwah, NJ: Lawrence Erlbaum Associates.
- Emmorey, Karen & Judy S. Reilly (1995). Theoretical issues relating language, gesture and space: An overview. In Karen Emmorey & Judy S. Reilly (Eds.), *Language, gesture and space* (pp. 1–16). Hillsdale, NJ: Lawrence Erlbaum Associates.
- Emmorey, Karen & Judy S. Reilly (1998). The development of quotation and reported action: Conveying perspective in ASL. In Eve V. Clark (Ed.), *Proceedings of the Twenty-Ninth Annual Child Research Forum* (pp. 81–90). Stanford, CA: CSLI Publications.
- Emmorey, Karen, Barbara Tversky, & Holly A. Taylor (2000). Using space to describe space: Perspective in speech, sign, and gesture. *Spatial Cognition and Computation*, 2, 157–180.
- Enfield, Nick J., Sotaro Kita, & Jan P. de Ruiter (2007). Primary and secondary pragmatic functions of pointing gestures. *Journal of Pragmatics*, 39, 1722–1741.
- Engberg-Pedersen, Elisabeth (1993). *Space in Danish Sign Language: The semantics and morpho-syntax of the use of space in a visual language*. Hamburg: Signum Press.
- Fadjo, Cameron L., Ming-Tsan P. Lu, & John B. Black (2009). Instructional embodiment and video game programming in an after-school program. Presented at *Ed-Media: World conference on education multimedia, hypermedia and telecommunications*, Honolulu, Hawaii.
- Fauconnier, Gilles & Mark Turner (2002). *The way we think: Conceptual blending and the mind's hidden complexities*. New York: Basic Books.

- Fox, Barbara A. & Jessica Robles (2010). It's like mmm: Enactments with *it's like*. *Discourse Studies*, 12, 715–738.
- Gerofsky, Susan (2010). Mathematical learning and gesture: Character viewpoint and observer viewpoint in students' gestured graphs of functions. *Gesture*, 10 (2/3), 321–343.
- Goldin-Meadow, Susan & David McNeill (1999). The role of gesture and mimetic representation in making language the province of speech. In Michael Corbalis & Stephen E. G. Lea (Eds.), *The descent of mind: Psychological perspective on hominid evolution* (pp. 155–172). Oxford: Oxford University Press.
- Haviland, John B. (1996). Projections, transpositions, and relativity. In John J. Gumperz & Stephen C. Levinson (Eds.), *Rethinking linguistic relativity* (pp. 271–323). Cambridge: Cambridge University Press.
- Hostetter, Autumn B. & Martha W. Alibali (2008). Visible embodiment: Gestures as simulated action. *Psychonomic Bulletin and Review*, 15 (3), 495–514.
- Hostetter, Autumn B. & Martha W. Alibali (2010). Language, gesture, action! A test of the gesture as simulated action framework. *Journal of Memory and Language*, 63, 245–257.
- Hutchins, Edwin (1996). *Cognition in the wild*. Cambridge, MA: Bradford Books.
- Janzen, Terry (2004). Space rotation, perspective shift, and verb morphology in ASL. *Cognitive Linguistics*, 15 (2), 149–174.
- Janzen, Terry (2012). Two ways of conceptualizing space: Motivating the use of static and rotating vantage point space in ASL discourse. In Barbara Dancygier & Eve Sweetser (Eds.), *Viewpoint in language: A multimodal perspective* (pp. 156–175). Cambridge: Cambridge University Press.
- Kendon, Adam (1980). Gesture and speech: Two aspects of the process of utterance. In Mary Ritchie Key (Ed.), *Nonverbal communication and language* (pp. 207–227). The Hague: Mouton.
- Kendon, Adam (2004). *Gesture: Visible action as utterance*. Cambridge: Cambridge University Press.
- 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.
- Liddell, Scott K. (2003). *Grammar, gesture and meaning in American Sign Language*. Cambridge: Cambridge University Press.
- Liddell, Scott K. & Melanie Metzger (1998). Gesture in sign language discourse. *Journal of Pragmatics*, 30 (6), 657–697.
- Lillo-Martin, Diane (1995). The point of view predicate in American Sign Language. In Karen Emmorey & Judy S. Reilly (Eds.), *Language, gesture, and space*. Hillsdale, NJ: Lawrence Erlbaum Associates.
- Marentette, Paula, Natasha Tuck, Elena Nicoladis, & Simone Pika (2004). The effects of language, culture and embodiment on signed stories. Paper presented at the *Theoretical Issues in Sign Language Research 8*, Barcelona, Spain. Retrieved from <http://www.ub.edu/ling/tislr8/Marentette-Tuck-Nicoladis-Pika.doc>, August 14, 2012.
- Maury-Rouan, Claire (2011). Voices and bodies: Investigating nonverbal parameters of the participation framework. In Gale Stam & Mike Ishino (Eds.), *Integrating gestures: The interdisciplinary nature of gesture* (pp. 309–320). Amsterdam: John Benjamins.
- McClave, Evelyn Z. (2000). Linguistic functions of head movements in the context of speech. *Journal of Pragmatics*, 32, 855–878.

- McCleary, Leland & Evani Viotti (2009). Sign-gesture symbiosis in Brazilian Sign Language narrative. In Fey Parrill, Vera Tobin, & Mark Turner (Eds.), *Conceptual structure, discourse and language* 9 (pp. 181–201). CSLI Publications.
- McNeill, David (1996). *Hand and mind: What gestures reveal about thought*. Chicago: University of Chicago Press.
- McNeill, David (2005). *Gesture and thought*. Chicago: University of Chicago Press.
- McNeill, David & Susan Duncan (2000). Growth points in thinking-for-speaking. In David McNeill (Ed.), *Language and thought* (pp. 141–161). Cambridge: Cambridge University Press.
- Merola, Giorgio (2009). The effects of the gesture viewpoint on the students' memory of words and stories. *Gesture-based Human-Computer Interaction and Simulation Lecture Notes in Computer Science*, 5085, 272–281.
- Metzger, Melanie (1995). Constructed dialogue and constructed action in American Sign Language. In Cecil Lucas (Ed.), *Sociolinguistics in deaf communities*. Washington, DC: Gallaudet University Press.
- Mittelberg, Irene (2006). Methodology for multimodality: One way of working with speech and gesture data. In Monica Gonzalez-Marquez, Irene Mittelberg, Seana Coulson, & Michael J. Spivey (Eds.), *Methods in cognitive linguistics* (pp. 225–248). Amsterdam: John Benjamins.
- Narayan, Shweta (2012). Maybe what it means is he actually got the spot: Physical and cognitive viewpoint in a gesture study. In Barbara Dancygier & Eve Sweetser (Eds.), *Viewpoint in language: A multimodal perspective* (pp. 113–135). Cambridge: Cambridge University Press.
- Núñez, Rafael & Eve 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 (3), 401–450.
- Ochs, Elinor, Sally Jacoby, & Patrick Gonzales (1994). Interpretive journeys: How physicists talk and travel through graphic space. *Configurations*, 2 (1), 151–171.
- Özyürek, Asli, Sotaro Kita, Shanley Allen, Reyhan Furman, & Amanda Brown (2005). How dose linguistic framing of events influence co-speech gestures? Insight from cross-linguistic variations and similarities. *Gesture*, 5 (1/2), 219–240.
- Özyürek, Asli & Pamela Perniss (2011). Event representation in sign language: A crosslinguistic perspective. In Jürgen Bohnemeyer & Eric Pederson (Eds.), *Event representation in language: Encoding events at the language-cognition interface* (pp. 84–107). Cambridge: Cambridge University Press.
- Padden, Carol (1990). The relation between space and grammar in ASL verb morphology. In Ceil Lucas (Ed.), *Sign language research: Theoretical issues* (pp. 118–132). Washington, DC: Gallaudet University Press.
- Park, Yujung (2009). Interaction between grammar and multimodal resources: Quoting different characters in Korean multiparty conversation. *Discourse Studies*, 11 (79), 79–104.
- Park-Doob, Mischa A. (2010). Gesturing through time: Holds and intermodal timing in the stream of speech. Unpublished doctoral dissertation, University of California, Berkeley.
- Parrill, Fey (2009). Dual viewpoint gestures. *Gesture*, 9 (3), 271–289.
- Parrill, Fey (2010a). The hands are part of the package: Gesture, common ground, and information packaging. In John Newman & Sally Rice (Eds.), *Empirical and experimental methods in cognitive/functional research* (pp. 285–302). Stanford: CSLI Publications.
- Parrill, Fey (2010b). Viewpoint in speech-gesture integration: Linguistic structure, discourse structure, and event structure. *Language and Cognitive Processes*, 25 (5), 650–668.

- Parrill, Fey (2012). Interactions between discourse status and viewpoint in co-speech gesture. In Barbara Dancygier & Eve Sweetser (Eds.), *Viewpoint in language: A multimodal perspective* (pp. 97–112). Cambridge: Cambridge University Press.
- Parrill, Fey, Jennifer Bullen, & Huston Hoburg (2010). Effects of input modality on speech-gesture integration. *Journal of Pragmatics*, 42 (11), 3130–3137.
- Parrill, Fey & Eve Sweetser (2004). What we mean by meaning: Conceptual integration in gesture analysis and transcription. *Gesture*, 4 (2), 197–219.
- Perniss, Pamela (2007a). *Space and iconicity in German Sign Language (DGS)*. (MPI Series in Psycholinguistics, 45). Nijmegen: Max Planck.
- Perniss, Pamela (2007b). Achieving spatial coherence in German Sign Language narratives: The use of classifiers and perspective. *Lingua*, 117 (7), 1315–1338.
- Perniss, Pamela & Asli Özyürek (2008). Representations of action, motion and location in sign space: A comparison of German (DGS) and Turkish (TID) Sign Language narratives. In Josep Quer (Ed.), *Signs of the time: Selected papers from TISLR 8* (pp. 353–378). Hamburg: Signum Press.
- Peters, Cynthia (2000). *Deaf American literature: From carnival to the canon*. Washington, DC: Gallaudet University Press.
- Petrack, Carmen, Matthew Berland, & Taylor Martin (2011). Allocentrism and computational thinking. *CSCL 2011 Proceedings*, vol. 2. Retrieved from: <http://www.activelearninglab.org/wp-content/uploads/2011/05/CSCL2011PetrackBerlandMartin1.pdf>, August 14, 2012.
- Pfau, Roland & Josep Quer (2010). Nonmanuals: Their grammatical and prosodic roles. In Diane Brentari (Ed.), *Sign languages*. Cambridge: Cambridge University Press.
- Quinto-Pozos, David (2007). Can constructed action be considered obligatory? *Lingua*, 117 (7), 1285–1314.
- Rayman, Jennifer (1999). Storytelling in the visual mode: A comparison of ASL and English. In Elizabeth Winston (Ed.), *Storytelling and conversation: Discourse in deaf communities* (pp. 59–82). Washington, DC: Gallaudet University Press.
- Schembri, Adam (2003). Rethinking “classifiers” in signed languages. In Karen Emmorey (Ed.), *Perspectives on classifier constructions in sign languages* (pp. 3–34). Hillsdale, NJ: Lawrence Erlbaum Associates.
- Schembri, Adam, Carloine Jone, & Denis Burnham (2005). Comparing action gestures and classifier verbs of motion: Evidence from Australian sign language, Taiwan sign language, and nonsigners’ gestures without speech. *Journal of Deaf Studies and Deaf Education*, 10 (3), 272–290.
- Shaffer, Barbara (2012). Reported speech as an evidentiality strategy in American sign language. In Barbara Dancygier & Eve Sweetser (Eds.), *Viewpoint in language: A multimodal perspective* (pp. 139–155). Cambridge: Cambridge University Press.
- Sidnell, Jack (2006). Coordinating gesture, talk and gaze in reenactments. *Research on Language and Social Interaction*, 39 (4), 377–409.
- Streeck, Jürgen (2002). Grammar, words and embodied meanings: On the evolution and uses of *so* and *like*. *Journal of Communication*, 52, 581–596.
- Sweetser, Eve (2006). Looking at space to study mental spaces: Co-speech gesture as a crucial data source in cognitive linguistics. In Monica Gonzalez-Marquez, Irene Mittelberg, Seana Coulson, & Michael J. Spivey (Eds.), *Methods in cognitive linguistics* (pp. 201–224). Amsterdam: John Benjamins.

- Sweetser, Eve (2012). Introduction: Viewpoint and perspective in language and gesture, from the ground down. In Barbara Dancygier & Eve Sweetser (Eds.), *Viewpoint in language: A multimodal perspective* (pp. 1–23). Cambridge: Cambridge University Press.
- Taub, Sarah F. (2001). *Language from the body: Iconicity and metaphor in American sign language*. Cambridge: Cambridge University Press.
- Verhagen, Arie (2005). *Constructions of intersubjectivity: Syntax, grammar and cognition*. Oxford: Oxford University Press.
- Wright, Tracey (2001). Karen in motion: The role of physical enactment in developing an understanding of distance, time, and speed. *Journal of Mathematical Behavior*, 20, 145–162.
- Young, Katharine (2002). The dream body in somatic psychology: The kinaesthetics of gesture. *Gesture*, 2 (1), 45–70.
- Zwitserlood, Inge (2003). *Classifying hand configurations in Nederlandse Gebarentaal (Sign Language of the Netherlands)*. Utrecht: LOT.

Author's address

Kashmiri Stec
Dept. of Communication and Information Sciences
University of Groningen
Oude Kijk in 't Jatstraat 26
9712 EK Groningen
The Netherlands
k.k.m.stec@rug.nl

About the author

Kashmiri Stec is a cognitive scientist who studies multimodal communication within the frameworks of embodied and situated cognition.