Published in final edited form as:
Spat Cogn Comput. 2015 June 1; 15(3): 143–169. doi:10.1080/13875868.2014.1003933.

Viewpoint in the Visual-Spatial Modality: The Coordination of Spatial Perspective

Jennie E. Pyers[a], **Pamela Perniss**[b], and **Karen Emmorey**[c]

[a]Wellesley College, Psychology Department, Wellesley, MA 02481, USA

[b]University of Brighton, School of Humanities, Checkland Building, BN1 9PH Brighton, UK, p.perniss@brighton.ac.uk

[c]San Diego State University, Laboratory for Language and Cognitive Neuroscience, 6495 Alvarado Road, Suite 200, San Diego, CA 92120, kemmorey@projects.sdsu.edu

Abstract

Sign languages express viewpoint-dependent spatial relations (e.g., left, right) iconically but must conventionalize from whose viewpoint the spatial relation is being described, the signer's or the perceiver's. In Experiment 1, ASL signers and sign-naïve gesturers expressed viewpoint-dependent relations egocentrically, but only signers successfully interpreted the descriptions non-egocentrically, suggesting that viewpoint convergence in the visual modality emerges with language conventionalization. In Experiment 2, we observed that the cost of adopting a non-egocentric viewpoint was greater for producers than for perceivers, suggesting that sign languages have converged on the most cognitively efficient means of expressing left-right spatial relations. We suggest that non-linguistic cognitive factors such as visual perspective-taking and motor embodiment may constrain viewpoint convergence in the visual-spatial modality.

Keywords

Spatial language; viewpoint; sign language; gesture

1. INTRODUCTION

Successful communication about spatial relations depends in large part on the speaker and listener converging on a common viewpoint on a spatial scene. Speakers around the world have conventionalized different ways to clearly convey spatial relations. Some languages use primarily an allocentric frame of reference that describes spatial relations using an absolute coordinate system such as North, South, East, and West, e.g., the tent is *west* of the picnic table (Levinson, 2003). Speakers of other languages (e.g., English and many Indo-European languages) tend to use a relative frame of reference to talk about spatial relations and to describe the location of an object relative to another object, e.g. the tent is to the *left of* the picnic table. Speakers of languages that use a relative frame of reference, as opposed to an allocentric frame of reference, face a viewpoint coordination problem—*left* depends on

Corresponding Author Tel: 781-283-3736, Fax: 781-283-3730, jpyers@wellesley.edu.

the location of the person who views the picnic table. To overcome this problem, speakers will often add additional information to clarify the viewpoint, e.g. specifying their own viewpoint, as in *my left*, or their interlocutor's viewpoint, as in *your right*. A long history of research on spatial language and cognition supports the primacy of the speaker's egocentric viewpoint (e.g., Clark, 1973; Levelt, 1989; Miller & Johnson-Laird, 1976; Piaget & Inhelder, 1956). An egocentric description is directly available because it represents the speaker's perception of a scene and, as such, requires little cognitive effort to produce. Indeed, several studies show that English speakers preferentially describe spatial relations egocentrically, from their own viewpoint (Levelt, 1989; Miller & Johnson-Laird, 1976).

An egocentric description of spatial relations, however, is by no means the default for speakers. Social, cognitive, communicative, and situational factors may lead speakers to take into account the viewpoint of their addressee when describing spatial relations and to express spatial relations from their addressee's point of view instead of their own, e.g. saying "on your right" instead of "on my left" (Galati, Michael, Mello, Greenauer, & Avraamides 2013; Herrmann, Bürkle, & Nirmaier 1987; Mainwaring, Tversky, Ohgishi, & Schiano, 2003; Schober, 1993, 1995; Tversky & Hard, 2009). Depending on the circumstances, speakers in Schober's (1993) study adopted a non-egocentric perspective anywhere from 29% to 81% of the time. For jointly viewed scenes, speakers were more likely to adopt their addressee's viewpoint when their addressee had a different viewpoint on the scene (i.e., offset by either 90° or 180°). In addition, use of the addressee's viewpoint was even higher when the speaker was asked to record a message for non-present addressees, who potentially bore a greater cognitive load because they would be unable to ask for clarification (Schober, 1993). Similarly, speakers of English as well as speakers of Japanese (considered to be an extremely polite culture) were more likely to use a non-egocentric perspective when the addressee was unfamiliar with a spatial array compared to when the addressee was familiar with the array (Mainwaring et al., 2003), suggesting that perspective use does not simply reflect considerations of politeness (Brown & Levinson, 1987). Thus, speakers seem to be sensitive to the addressee's cognitive effort in understanding spatial descriptions and adapt their utterances accordingly to ensure successful communication.

In addition to the addressee's cognitive effort, the speaker's own cognitive effort may influence perspective choice. For example, Galati and Avraamides (2013) found that speakers who sat at an oblique offset of 135° with respect to their present addressee were more likely to give spatial descriptions from an egocentric perspective. In this situation, representing a non-egocentric perspective involved a more difficult spatial computation. When the cognitive effort was minimized (as when speakers sat at 90° or 180° relative to their addressee), speakers were more likely to give descriptions from their addressee's perspective, confirming Schober's (1993) findings.

Overall, the preference for an egocentric or non-egocentric perspective in these situations may be explained in terms of the relative cognitive burdens placed on speaker and addressee. Mainwaring et al. (2003) suggest that the transformation of a spatial scene into a linguistic utterance (scene-to-language) or the transformation of the utterance into a spatial image (language-to-scene) dictates the relative cognitive burden of viewpoint alignment for speaker and addressee. When comprehending a spoken spatial description, the addressee must mentally map the language input onto the spatial scene, and this mapping is more difficult if the addressee must interpret the spatial relations from a viewpoint different from their own. Mainwaring et al. (2003) further suggest that transforming a spoken utterance into a mental image of a spatial situation is more difficult than generating language to describe a known spatial scene, regardless of the perspective – egocentric or non-egocentric. When speakers adopt the (non-shared) perspective of their addressee during their spatial descriptions, they remove the more difficult perspective-taking demands from the addressee's mapping task, presumably leading to more rapid comprehension and more successful communication. In general, the choice of non-egocentric over egocentric spatial descriptions reflects the need for perspective alignment in successful joint communication, and the more salient this need is, the more likely it is that the speaker will adopt the perspective of their listener.

To date, research on viewpoint in spatial descriptions has primarily focused on spoken rather than signed languages, leaving unanswered the nature of viewpoint convergence in the visual-spatial modality. Crucially, the cognitive requirements that constrain the convergence of viewpoint in spatial descriptions may vary as a function of language modality. Sign languages differ from spoken languages in how spatial relations are linguistically marked. Instead of using spatial terms like “right” or “left,” signers of many different and unrelated sign languages directly map spatial relations onto the signing space in front of the body. That is, signers create an iconic representation in which real-world spatial relationships are mapped onto the relative positions of their hands. In essence, the signer's hands recreate the target spatial relationship in signing space. This recreation can be done from the signer's own perspective of a real or imagined scene as illustrated in Figure 1A. To describe a room containing a table on the left, the signer can produce a classifier sign for the round table to his left to indicate that the table is on the left. As shown in Figure 1B, the signer could also produce the classifier sign for the table on his right so that the table is on the left of signing space as viewed by an addressee who is facing him (just as you, the reader, are facing the signer in this illustration).

Regardless of the signer's choice of viewpoint, the perceiver faces a unique challenge in interpreting iconic descriptions of left-right spatial relations. The canonical position of signing interlocutors is face-to-face, which means that a signer and addressee view spatial descriptions from opposite positions: what a signer articulates on the left side of signing space is perceived by the addressee on his right, and vice versa, as illustrated in Figure 1. If signers communicate left-right spatial relations from an egocentric viewpoint (e.g., on the left from the signer's perspective) and perceivers interpret them from their own egocentric viewpoint (e.g., on the right from the perceiver's perspective), then the accurate expression and communication of the spatial relationship would fail. While the analogical representation of spatial relations seems relatively transparent, it is only when viewpoint expression and interpretation are conventionalized that signed spatial descriptions can be successful.

Such coordination of perspectives in sign languages only needs to take place when talking about non-present referents or present referents that are not jointly viewed, i.e. for “non-shared” space (see Figure 2a). For present referents that are visible to both the signer and addressee, the spatial mapping remains the same from any perspective because the hands map directly onto the objects in the shared environment, as illustrated in Figure 2b (Emmorey & Tversky, 2002). In the current study, we are concerned only with spatial descriptions of scenes that are not jointly viewed by the interlocutors, and that thus require perspective coordination. Under these circumstances, successful communication depends on one of the two communication partners taking the other's perspective: either the signer conveys the left-right spatial relation from the perceiver's perspective so the perceiver can interpret the description from her own egocentric perspective (Figure 1B), or the signer describes the spatial relation egocentrically and the perceiver must interpret the spatial description non-egocentrically, i.e., from the signer's perspective (Figures 1A and 2A).

Across unrelated sign languages in which spatial relations are depicted iconically by the placement of classifier signs in space (as in Figure 1), signers have been found to produce such descriptions from an egocentric viewpoint (Emmorey, 1996; Emmorey, Klima, & Hickok, 1998; Perniss, 2007; Pyers, Perniss, & Emmorey, 2008). Pyers et al. (2008) asked pairs of signers from eight different sign languages[1] to engage in a picture description task, in which the signers described pictures containing spatial scenes to an addressee, who selected the target picture from an array (the Man and Tree task, Space Stimuli Kit 1.1; Max Planck Institute for Psycholinguistics Field Manual, 1992). Signers across all eight sign languages produced egocentric left-right spatial descriptions (only rarely providing left-right descriptions from their addressee's viewpoint), and addressees overwhelmingly picked the correct picture, interpreting the description from the signer's egocentric viewpoint. These data indicate the conventionalized nature of this viewpoint strategy across many sign languages. The preference for egocentric descriptions across sign languages stands in contrast to speakers’ general willingness to adopt a non-egocentric viewpoint to support their addressees’ comprehension under similar circumstances. In sign languages, in contrast to spoken languages, the addressee is the one who relinquishes his/her viewpoint.

The consistency observed in viewpoint choice within signers and across distinct sign languages suggests that there are properties of representing spatial relations analogically in the visual-spatial modality that favor representing viewpoint from the signer's perspective, even when the need for perspective alignment is salient. For example, within such an analogical system, the egocentric representation of left-right spatial relations involves a direct mapping of the spatial relation onto the left-right position of the hands, as in Figure 1A, whereas the adoption of the non-egocentric addressee's perspective would require that the signer override his/her mental image of the scene to produce the opposite visual-manual description (e.g., placing the sign for the table on his right in the signing space, even though the table was on his left in the perceived scene, Figure 1B). In contrast, for spoken spatial descriptions there is no potential for a perceptual conflict between the linguistic expression and the observed spatial scene. The signer's task in producing a non-egocentric spatial description may thus be more cognitively demanding than what is required for the non-egocentric production of arbitrary spoken lexical labels. Further, the addressee's ability to

[1]American Sign Language, Deutsche Gebärdensprache (German Sign Language), Indo-Pakistani Sign Language, Israeli Sign Language, Lingua Italiana dei Segni (Italian Sign Language), Nederlandse Gebarentaal (Sign Language of the Netherlands), Nihon Shuwa (Japanese Sign Language), and Russian Sign Language.

interpret a spatial description from the signer's perspective may be facilitated by the ability to engage in motor embodiment (Emmorey et al., 1998). That is, sign perception might involve an internal simulation of sign production, as if the sign perceiver were signing herself (Emmorey, 2006). Thus, the relative cognitive burdens placed on the producer and perceiver with respect to viewpoint convergence may be differently distributed within an iconic, analog system of spatial representation.

In the current study, we investigated whether the iconically driven property of using space to describe spatial relations constrains viewpoint choice in the visual-spatial modality. In Experiment 1, we directly compared the production of spatial descriptions in ASL by Deaf fluent signers to the use of gesture (without speech) by sign-naïve English speakers unaccustomed to using the visual-spatial modality as the sole channel of communication. Experiment 1 allowed us first to replicate previous findings that ASL signers favor an egocentric expression and non-egocentric interpretation of viewpoint-dependent spatial relations. In addition, by directly comparing ASL signers to gesturers on a similar task, we could investigate whether it is indeed properties of the visual modality that constrain the representation and coordination of spatial viewpoint.

We hypothesized that the iconic, analogical expression of spatial relations in which the relation between objects is recreated by the placement of the hands in space leads to a strong preference for an egocentric representation of left-right spatial relations, and as such, both signers and gesturers should express spatial relations from their own point of view. We further hypothesized that if the pattern of viewpoint alignment observed in sign language is driven by properties of the visual-spatial modality, then non-signers should interpret manual expressions of spatial relations non-egocentrically. Such a pattern in the gesturers would be quite revealing about the effect of modality on viewpoint alignment as these native English speakers would override their spoken language bias to give non-egocentric addressee-viewpoint descriptions in this type of communication task. Moreover, English-speaking addressees accustomed to interpreting spatial descriptions from their own viewpoint would instead have to take on the demands of perspective-taking themselves and interpret the utterance from a non-egocentric viewpoint.

In Experiment 2, we assessed the relative cognitive costs of producing and comprehending left-right spatial relations in the visual-spatial modality by randomly assigning pairs of gesturers to conditions in which they were instructed to produce and comprehend spatial descriptions either egocentrically or non-egocentrically (or they were not assigned a specific perspective, as in Experiment 1). We hypothesized that there would be an asymmetry in the relative cognitive burdens of production and comprehension of gestured viewpoint-dependent spatial descriptions such that there would be a greater egocentric bias for production than for comprehension in the visual-spatial modality.

2. EXPERIMENT 1

2.1. Method

2.1.1. Participants—To replicate the previous research that ASL signers prefer an egocentric description of spatial relations, we recruited eight Deaf fluent ASL signers (6 females; M_{age}=35.2 *SD*=9.8) to participate in a communication task. To test our hypothesis that the preference for egocentric description exhibited in sign languages is inherent to communication of spatial relations in the visual-spatial modality, we also recruited eighteen female undergraduate non-signers (M_{age}=19.6, *SD*=1.2) to participate in the same communication task. The Deaf participants were all paid for their participation; the non-signers were either paid or received course credit for their participation. All participants were tested in pairs, with each randomly assigned the role of Producer or Perceiver.

2.1.2. Materials and Procedure—Pairs were seated facing each other separated by an eleven-inch divider, but always in view of each other. We provided the Producer with a binder of 16 pictures including 8 pictures that depicted a man standing to the left or right of a tree, and 8 filler items depicting a man in front of or behind a tree (Figure 3).[2] The set of pictures we used were adapted from The Man and Tree task, (Space Stimuli Kit 1.1; Max Planck Institute for Psycholinguistics Field Manual, 1992) by Senghas (2010a). Producers viewed each picture one at a time in one of two fixed random orders; each picture was separated by a blank piece of paper. The Perceivers' pictures were laid out on the table in a random array. The divider between participant pairs prevented the Producer and Perceiver from jointly viewing the spatial scene being described (i.e., the stimulus picture). In contrast to previous studies with spoken languages, this kind of shielded communication task was necessary because when signers and perceivers jointly view a scene, they use shared space in their signed descriptions (see Figure 2b). For spoken languages, the opposite is the case: jointly-viewed, real object arrays elicit viewpoint marking, but shielded communication tasks using photographs of object arrays do not.

ASL signers completed the task in ASL, and non-signers completed the task using only two gestures that they were taught, one to represent a man, and one to represent a tree (Figure 4). Because our research question focused on the representation of left-right spatial relations, participants were neither explicitly instructed to represent the orientation of the man, nor told which side of the "man" handshape represented the front or back of the body.[3] All Producers were instructed to look at the picture in front of them, remember the configuration, turn the page, and produce a signed or gestured description of the picture that they viewed. Gesture Producers were instructed not to speak during the task. Perceivers were instructed to select the picture that best matched the Producer's description, show it to the experimenter (keeping it out of the Producer's view), and then return it to the array. They could ask the Producer to repeat their spatial description. They were additionally told that they could select the same picture for more than one item; this instruction kept Perceivers' selections from being influenced by the process of elimination. Because we were interested in the patterns of viewpoint expression and comprehension in the visual modality, we allowed participants to take their time with their productions and selections in order for them to best convey the information in the picture and to make the most accurate selection that

[2]Although front/back spatial relations require perspective alignment, we focus only on left/right relations here. In contrast to left/right, front/back specifications are less conventionalized in ASL: they are less consistently marked and less consistently associated with an egocentric perspective (Emmorey 1996, 2001).

[3]The handshape used to convey the man (see Figure 4a) does not have a strongly conventionalized front or back intrinsic side in ASL. In some cases, the back of the hand can represent the front of the body, as when the handshape is used in the sign WALK, but in other cases the back of the hand represents the back of the body, as when the handshape is used in LIE-DOWN.

matched the gestural configuration. Participants began with three practice items depicting two trees or two men next to each other, and then proceeded to the test items. The three practice items were symmetrical and did not require viewpoint marking. Although both ASL signers and non-signers completed the practice items, the practice items were present primarily to offer the non-signers an opportunity to practice producing the newly learned gestures. All participants were videotaped for coding offline.

2.1.3. Coding—The eight items that required viewpoint marking to distinguish left-right spatial relations were coded. We coded the accuracy of the Perceivers' selections by scoring as successful any selection that correctly mapped the left-right relation. Thus, although Producers may have encoded the orientation of the man with respect to the tree by orienting the hand in a corresponding way, our coding focused only on the left-right mapping.

We coded whether Producers expressed the spatial relation from their own viewpoint or from the Perceiver's viewpoint (Figure 5). For example, in viewing the picture of the man standing to the left of the tree, the signer could directly map the locations of the signs for the tree and the man onto the same left-right locations in signing space, matching his/her view of the picture. Such a strategy was classified as the Producer using his/her own "Producer's Perspective." Alternatively, the Producer could reverse the locations of the signs and describe the spatial relation from the Perceiver's perspective, such that the signs would directly map to the pictured objects from the Perceiver's viewpoint. This strategy was coded as "Perceiver's Perspective." Every description produced by the Producer was coded as an individual utterance. For example, the Producer may have repeated a description of a single picture more than once if the Perceiver asked for clarification. If the signer repeated a description and adopted a different viewpoint strategy, one of their descriptions would be coded as "Producer's Perspective" and the second as "Perceiver's Perspective." Thus, for each Producer, the percentage of spatial descriptions produced using each viewpoint was computed.

Similarly, we coded whether Perceivers interpreted the description of the spatial relation from their own viewpoint or from the Producer's viewpoint. For example, the Producer could express a spatial relation by placing the sign for 'man' on his/her left side of signing space, and the sign for 'tree' on his/her right side of signing space. If the Perceiver interpreted this relationship as the man standing to the left of the tree, as evidenced by the Perceiver's picture selection, the interpretation was coded as being from the "Producer's Perspective." If, on the other hand, the Perceiver interpreted the description as displaying a man standing to the right of the tree, it was coded as "Perceiver's Perspective", as this mapped onto the Perceiver's own view of the signs.

2.2. Results

All four of the ASL Producers used their own viewpoint 100% of the time, and the ASL Perceivers readily adopted the Producer's Perspective 91% of the time (See Table 1); this viewpoint compatibility corresponded to successful selections on 91% of trials. The small amount of variability in the Perceivers' performance reflected the fact that three of the four Perceivers each selected a picture representing the Perceiver's Perspective once.

Like signers, gesture Producers conveyed the spatial relations predominantly from their own point of view, adopting an egocentric perspective at a rate that did not differ significantly from the ASL signers ($t(11)=1.25$, $p=.24$; see Table 1). Gesturers were quite variable as a group: five Producers expressed egocentric viewpoints exclusively, two did so more than 75% of the time, while the remaining two participants adopted the non-egocentric Producer's Perspective more than 75% of the time, although neither did so exclusively. Thus, while there was no statistical difference between the ASL and gesture groups, there was a qualitative difference in consistency, albeit with a strong preference for the egocentric viewpoint. Unlike the ASL Perceivers, however, the gesture Perceivers rarely interpreted the Producers' gestures as reflecting the Producers' perspective; only four of the nine participants adopted the non-egocentric perspective, and each did so for only one trial. Gesture Perceivers chose their own egocentric viewpoint for interpretation on the majority of trials and significantly more often than what was observed for the ASL signers ($t(11)=2.47$, $p<.0001$; see Table 1). This pattern in the gesturers resulted in a viewpoint incompatibility – with neither participant relinquishing their egocentric perspective – in the majority of trials, leading to only 17 correct matches out of 72 in the entire sample for the eight target items. The viewpoint compatibility necessary for correct matches and successful communication was thus only infrequently achieved. The majority of successful matches occurred when Producers communicated the spatial relation from the Perceiver's Perspective, taking a non-egocentric perspective, while the Perceiver maintained an egocentric perspective in interpretation (see Figure 6).

2.3. Discussion

In Experiment 1, even with a small sample size, we robustly replicated previous work on sign languages demonstrating that signers prefer to express and interpret viewpoint-dependent spatial relations from the signer's egocentric perspective (Emmorey & Tversky, 2002; Pyers et al., 2008). In addition, we found that sign-naïve gesturers expressed viewpoint-dependent spatial relations from an egocentric perspective, indicating the immediate availability of this perspective. In contrast to English speakers, who under a similar face-to-face paired communication task readily adopt a non-egocentric perspective to convey viewpoint-dependent spatial relations for their addressees (Schober, 1993), both ASL signers and sign-naïve gesturers almost exclusively expressed spatial relations from their own viewpoint, rarely adopting the perspective of their addressee. Further, this strategy was understood by ASL Perceivers (but not by gesture Perceivers), who predominantly adopted a non-egocentric perspective to correctly interpret the descriptions of viewpoint-dependent spatial relations. Although this paradigm did not elicit non-egocentric productions on the part of signers, there may be other discourse situations, as has been observed with speakers, that might elicit non-egocentric productions. In this face-to-face communication task, however, we clearly see the conventionalization of viewpoint in ASL spatial descriptions.

We hypothesized that the consistent pattern of viewpoint representation and interpretation that has been observed across multiple sign languages may emerge because of the iconic, analog properties of the visual-spatial modality. We found that although sign-naïve gesturers egocentrically expressed left-right spatial relations in the visual-spatial modality, their communication partners did not consistently interpret the spatial expression as from the Producer's point of view. That the gesturers did not exhibit the same degree of viewpoint alignment observed for signers suggests that viewpoint conventionalization is a necessary process for visual-spatial languages, and one that must emerge over time.

Evidence from an emerging sign language, Nicaraguan Sign Language (NSL), provides support for this idea. The current form of NSL does not yet exhibit a systematic means of representing and interpreting spatial viewpoint (Pyers, Shusterman, Senghas, Spelke, & Emmorey, 2010; Senghas, 2010b). NSL signers sometimes convey the viewpoint from their own perspective and sometimes from their addressee's perspective. The variable performance of the Nicaraguan signers shows that conventionalization of viewpoint does not emerge early in the development of a sign language. However, one reason for the late emergence of viewpoint conventionalization may be related to the cognitive abilities of this population. The same signers who inconsistently represent viewpoint also show delays on theory-of-mind tasks that require understanding that other people have different perspectives (Pyers & Senghas, 2009). Thus, the lack of conventionalization observed in the language may be related to a poor understanding of what other people know (and don't know) and of the kinds of information that are relevant for successful communication. These pragmatic features related to theory-of-mind development could drive conventionalization in a language. The gesturers in our study, on the other hand, were typically developing undergraduates with presumably mature theory-of-mind abilities, but with limited experience communicating solely in the visual-spatial modality. As such, the gesturers were not experienced enough with the visual-spatial modality to understand the need for perspective convergence.

That sign languages need to conventionalize perspective for the accurate communication of viewpoint-dependent spatial relations does not resolve the question of why sign languages seem to converge on a pattern of expressing and interpreting viewpoint from the signer's rather than the addressee's perspective, as is observed in spoken languages. In Experiment 2, we tested our hypothesis that the pattern of viewpoint convergence in sign languages is a direct result of the asymmetry in cognitive cost for the signer and the addressee when adopting a non-egocentric perspective.

For spoken language, adopting a non-egocentric perspective incurs a cognitive cost for the speaker and likely also for the addressee. It is unknown, however, whether this cognitive cost is greater for the speaker compared to the addressee. We hypothesized that in the visual modality the cognitive cost for the signer is greater than for the addressee because of the iconic nature of spatial language. As noted in the introduction, for the signer to adopt the addressee's perspective, he or she must generate a spatial description that is in direct perceptual (visual and somatosensory) conflict with the spatial scene. We suggest that this perceptual conflict presents a relatively greater cognitive load for the signer compared with the addressee, whose cognitive load may be reduced by the ability to internally simulate the signer's manual productions (i.e., motor embodiment; see Pickering and Garrod, 2013). Thus, in Experiment 2, we predicted that the cognitive cost to adopt a non-egocentric perspective would be greater for the gesture Producer than for the Perceiver.

3. EXPERIMENT 2

3.1. Method

3.1.1. Participants—We randomly assigned 94 female sign-naïve undergraduates (M_{age}=19.6 years; range, 18-23 years) to one of three conditions: Unassigned Perspective (*N*=22), Producer's Perspective (*N*=34), or Perceiver's Perspective (*N*=38).[4] All participants were tested in pairs, and within each pair, participants were randomly assigned the role of Producer or of Perceiver.

3.1.2. Materials and Procedures—We followed the same procedure as in Experiment 1, but the stimuli included only objects that had no intrinsic front/back features (a cup, a piece of paper, a chopstick, and a ball). The 12 test items depicted two of the four objects next to each other (Figure 7). The pictures were presented to the Producer in a binder in one of three random orders. The Perceiver's pictures were laid out in a fixed array that was consistent across all subjects (Figure 7), and the Producer viewed the Perceiver's array of pictures to better understand the Perceiver's possible choices. Producers and Perceivers were taught one-handed gestures for the objects (Figure 8), and they practiced producing and perceiving the gestures during four practice items that depicted the individual objects.

In the Unassigned Perspective condition, Producers and Perceivers received the same instructions as participants in Experiment 1. In the Producer's Perspective condition, participants were explicitly told to gesture and interpret the spatial relation from the Producer's perspective; with these instructions, the Producer was required to use an egocentric perspective while the Perceiver adopted the non-egocentric perspective of the Producer. In the Perceiver's Perspective condition, participants were told to gesture and to interpret the spatial relation from the perspective of the Perceiver. Here, the Producer relinquished her own perspective and produced a gesture configuration that was the opposite of what she saw, but that aligned with the Perceiver's view of the picture. Participants in the assigned perspective conditions were instructed to produce the gestures and to make their picture selections as fast as they could. All participants were videotaped, and the videotapes were digitized at 30fps for coding offline.

3.1.3. Coding—We coded the expression and interpretation of viewpoint in the same way as Experiment 1, and we determined the accuracy of the viewpoint expression for the two experimental conditions by identifying whether the participant correctly produced or interpreted the gesture using the assigned viewpoint. For the experimental conditions where perspectives were assigned, we also recorded how quickly the Producer made the correct gesture; we measured the duration from the time that they turned the stimulus page to the onset of the correct gestural configuration. We recorded how long it took the Perceiver to select the correct picture from the array by recording the duration of time that spanned from the onset of the Producer's gestural configuration to when the Perceiver touched the correct picture in the array. We used the time of gesture onset rather than offset because the gestures

[4]Due to experimenter error, one pair that had been pre-assigned to the Producer's Perspective condition actually received instructions for the Perceiver's Perspective condition. Thus, instead of having an equivalent number in each condition, there were 17 pairs in the Perceiver's Perspective condition and 19 pairs in the Producer's Perspective condition.

representing the objects and their relative positions could be readily perceived well before the offset of the gestural configuration. In addition, we observed some perceivers making their selections before the offset of the gesture. We excluded long reaction times by trimming the timing data to ±2SD from the mean time to produce the gesture or to ±2SD from the mean time to select the corresponding picture observed across all three conditions. This trimming reduced the number of correct items included in the Producer's timing analysis by 19 leaving a total of 370 correct items produced by 34 participants across both experimental conditions for the Producer's timing analysis. For the Perceivers, 8 correct items were trimmed, leaving a total of 351 correct items selected by 34 participants across both experimental conditions for the Perceiver's timing analysis. All items were included in analyses of viewpoint accuracy, and all items where the Producer conveyed the correct viewpoint were included in our measure of the Perceiver's picture selection accuracy.

3.2. Results

In the Unassigned Perspective condition we found an egocentric bias on the part of both Producers and Perceivers, just as we observed in Experiment 1. The Producers conveyed the spatial relations predominantly from their own point of view, with eight of the eleven Producers doing so for more than 75% of the trials (see Table 1). The Perceivers interpreted the gestures from their own egocentric viewpoint for the majority of trials, leading to only 57 out of 132 successful picture selections (see Table 1). Here, the majority of the successful picture matches (n=33) were made by three Perceivers (out of 11) who spontaneously adopted the (non-egocentric) perspective of the Producer for more than 75% of the trials.

We then compared performance in the Producer's Perspective condition to that in the Perceiver's Perspective condition. As shown in Table 2, Producers were significantly more accurate in their viewpoint production when they communicated the spatial relation from their own perspective than from the Perceiver's perspective. One Producer failed to adopt the Perceiver's perspective on all trials in that condition, but no such global perspective failures were observed in the egocentric Producer's perspective condition. The difference in the Producers' accuracy between conditions remained significant even when we eliminated the one Producer who failed to adopt the Perceiver's perspective when instructed, $t(17.39)=2.21$, $p=.04$.[5]

The advantage in the Producer's Perspective condition was also reflected in the timing of correct gesture production in the first trial, but not in the average gesture time across all trials (see Table 2): Producers who were required to gesture from a non-egocentric perspective were initially slower to produce their gestures than those allowed to communicate from an egocentric perspective, but they eventually adapted to the task such that the cost was no longer reflected in the aggregate timing across all trials.

For Perceivers, there was no significant difference in viewpoint accuracy between perspective conditions (see Table 2), and one Perceiver in each condition globally failed to adopt the correct viewpoint for all trials. There was no significant difference between groups in the number of correct picture selections made by Perceivers when the Producer conveyed

[5]When a Levene's test detects unequal sample variances, we report the corrected *t*-test, including fractional degrees of freedom.

the correct perspective (Producer's Perspective: *M*=9.58, *SD*=4.30; Perceiver's Perspective: *M*=8.63, *SD*=4.89), *t*(34)=.62, *p*=.53. We also observed no significant difference in how quickly Perceivers made a selection when they were required to comprehend descriptions from the Producer's perspective compared to their own perspective either on the first trial alone or across all 12 trials (see Table 2).

4. GENERAL DISCUSSION

Piaget and Inhelder (1956) were among the first to argue that perspective taking is challenging, and that a hallmark of mature cognition is the ability to take the perspective of another person. Yet several studies have shown that even neurotypical adults sometimes struggle with the additional mental computation required to take another perspective into account, making their online use of perspective-taking sometimes unreliable (Keysar, Barr, Balin, & Brauner, 2000; Keysar, Lin, & Barr, 2003). Successful communication and comprehension of viewpoint-dependent spatial relations, however, requires that one communication partner take on the additional cognitive load of adopting a non-egocentric perspective.

Viewpoint alignment does not occur spontaneously when communicating in the visual-spatial modality. In Experiment 1 and in the Unassigned Perspective condition of Experiment 2, gesturers who had limited experience with communicating in the visual-spatial modality did not automatically align their perspective to successfully describe viewpoint-dependent spatial relations via gesture. Although speakers have experience communicating in the visual spatial-modality using co-speech gestures, they may not be driven to perspective convergence in this modality because they can rely on information transmitted through speech. Little is known about viewpoint encoding and interpretation in co-speech gesture; however, Gullberg and Kita (2009: 258) mention in a footnote that addressees interpret directional information present in co-speech gesture egocentrically, just as we have observed in our current study. The degree to which speakers align viewpoint for spatial information communicated in co-speech gesture is an open question. When gesture is not accompanied by speech, however, viewpoint alignment is not automatic.

The lack of viewpoint alignment observed in the sign-naïve gesturers indicates that over time, viewpoint alignment may need to emerge in visual-spatial communication systems such as sign languages. However, it remains unclear what shapes the pattern of viewpoint alignment in sign languages. For example, only signers of NSL who exhibit a mature understanding of others' perspectives seem to have converged on a systematic way of expressing left-right relations (Pyers et al., 2010). The variability in the gesturers' performance in Experiment 1 and the Unassigned Perspective condition of Experiment 2 suggest that there may be individual differences in perspective-taking skill. In contrast, the homogeneity of the signing participants suggests that individual variability does not affect processes that have become conventionalized in the language, although see Quinto-Pozos, Singleton, Hauser, Levine, Garberoglio, and Hou (2013) for a discussion of how cognitive impairments in perspective-taking can negatively impact viewpoint understanding in signers. The current study cannot speak to the amount and kind of communicative experience in the visual-spatial modality that cognitively mature adults need in order to converge on a system for successful communication about viewpoint-dependent spatial relations. Perhaps if sign-naïve gesturers were communicating to an absent addressee (e.g. by means of a video recording that would later be shown to an addressee) we would see an increase in non-egocentric productions, as Schober (1993) observed with spoken languages. Further, if Producers and Perceivers were allowed to communicate about viewpoint during the task, we might see rapid viewpoint alignment because such interactions could highlight what aspects of the task are ambiguous.

The results from the ASL signers in Experiment 1 confirm what has been observed across many sign languages: the preferred pattern of viewpoint representation and interpretation is for the signer to maintain his/her own perspective and the interlocutor to adopt the signer's perspective. The consistency in viewpoint representation and interpretation suggests that this pattern may be driven by the iconic properties of the visual-spatial modality. Whereas English speakers may readily adopt another viewpoint on a spatial scene through a shift in possessive pronoun and change in spatial locative (*my left* to *your right*), such a viewpoint shift may be significantly more difficult in the visual modality. The results from Experiment 2 show that Producers provide more accurate spatial descriptions when instructed to convey the relation from their own point of view rather than from their interlocutor's point of view. Perceivers interpreted the viewpoint equally well when instructed to maintain an egocentric point of view or to adopt a non-egocentric point of view. Thus, the cost of adopting a non-egocentric viewpoint seems to be greater for Producers than for Perceivers, and as such sign languages appear to have converged on the most cognitively efficient means of expressing left-right spatial relations. At first, Producers were slower to initiate the gestured description from the Perceiver's perspective relative to their own perspective, but over multiple trials, this difference in response time between the two perspective conditions disappeared. The lack of overall difference in response time may have occurred because we used an offline task and gross measurements of time. Assessing processing time with more precise measures may provide clearer evidence for the online cognitive costs associated with adopting a non-egocentric versus an egocentric viewpoint in the visual-spatial modality.

The current study cannot identify the precise nature of the cognitive challenge, but we speculate that for Producers, taking a non-egocentric perspective in the visual-spatial modality may make additional demands on the executive function system, specifically conflict control, that do not arise for spoken language users. Conflict control involves inhibiting a prepotent response and performing an opposite response (e.g. saying “left” when looking at an arrow pointing right, or “night” when looking at a picture of the sun). Although children start to exhibit this ability at the end of the preschool years (Carlson, Mandell, & Williams, 2004), adults still find more complex versions of conflict control tasks cognitively taxing, as evidenced, for example, in the Stroop effect (Stroop, 1935) and, related to location, in the spatial Stroop and Simon effects (e.g. Lu & Proctor, 1995; MacLeod, 1991 for reviews). We speculate that to take on the Perceiver's perspective, signers face a conflict inhibitory control task that requires them to inhibit their own view or mental image of the left-right relation in a scene and produce the opposite, but nevertheless iconic, relation in signing space—a task that could be made even more difficult by the visual and somatosensory feedback they receive from their own signing. Because spoken languages do not represent spatial relations in an iconic, analog way, speakers do not face the same conflict between their image of a spatial scene and the vocal linguistic utterance that they produce. Critically, in contrast to signers, there is no conflicting visual or relevant somatosensory feedback. Without a direct conflict between the spatial image, perceived or imagined, and the linguistic utterance, speakers may find it much easier than signers to adopt the addressee's perspective, as is evidenced in the relatively higher rates of non-egocentric perspective use by English speakers in similar face-to-face communication tasks (e.g. Schober, 1993). Indeed, if sign languages relied primarily on lexical, non-analogical linguistic representations of space, we might predict a pattern of viewpoint alignment that more closely resembles what has been observed with English speakers, with Producers taking on the responsibility of representing spatial relations non-egocentrically. In terms of the relative cognitive burdens placed on communication partners, Mainwaring et al.'s (2003) idea that transforming a spatial scene into language is easier than transforming language into a spatial scene may hold primarily for spoken language viewpoint-dependent spatial descriptions. In the visual modality, transforming a perceived spatial scene into language for a communication partner may incur an additional cognitive cost that is not present for the spoken modality.

The Perceivers in Experiment 2 showed no greater cost when adopting the perspective of the Producer than when they were allowed to maintain their own egocentric viewpoint. We speculate that the Perceiver may recruit non-linguistic strategies to help them adopt the signer's perspective. Visual perspective-taking (i.e., putting yourself in someone else's shoes) is an easier task than mentally rotating objects in space (Amorim, Isableu, & Jarraya, 2006; Wraga, Creem, & Proffitt, 1999). Motor embodiment – the ability to imagine our bodies in another location in order to envision the view from that vantage point – seems to underpin spatial perspective-taking in non-linguistic tasks (Kessler & Thomson, 2010). Further, humans seem to rely on mentally transforming their own bodies to interpret information about another body (see Wilson, 2001, for discussion). In the case of sign languages, the perceivers' perspective-taking is likely facilitated by the fact that they can use motor simulation to mentally imitate the signer's manual productions, an argument that has been similarly made with respect to gesture comprehension (Alibali & Hostetter, 2010). By mapping their own body onto that of the signer, addressees could generate the correct viewpoint to accurately interpret the spatial relationship. However, this kind of embodiment is not available to sign producers when expressing a non-egocentric perspective; both looking at the spatial relation represented in the photograph from one's own perspective or imagining how a communication partner would perceive the same relation in the corresponding photograph would lead to the production of identical linguistic forms. Thus, internal simulations may only benefit sign language perceivers.

In essence, we suggest that analogical mapping of spatial relations in sign languages places different cognitive challenges on signers and addressees compared to the cognitive challenges faced by users of a spoken language where viewpoint-dependent spatial relations are primarily expressed with arbitrary words. The perceptual conflict encountered by the signer in producing a non-egocentric spatial description presents a clear cognitive load compared to the immediately available and non-conflicting egocentric perspective. There appears to be a discernible benefit to the signer in maintaining an egocentric perspective in spatial description. The cognitive burden on the signer could alone explain the cross-linguistic similarity in viewpoint alignment observed across different sign languages. For the perceiver, there seems to be no significant benefit to maintaining an egocentric perspective in interpretation. Therefore, the perceiver may more easily relinquish his/her egocentric perspective and can perhaps leverage non-linguistic cognitive tools such as motor embodiment to adopt a non-egocentric perspective.

This asymmetry in cognitive load for signer and addressee has implications for models of language processing with respect to whether and how audience design fits into the processing model. One suggestion is that the addressee's perspective and knowledge are already taken into account during the stage of message formulation, following a “principle of optimal design” from the outset (Clark, 1992; Clark & Murphy, 1982; Horton & Keysar, 1996). Another proposal holds instead that message planning is egocentric from the outset, and that the addressee's perspective is taken into account as needed, with monitoring for the needs of the addressee happening at later stages of production (Dell & Brown, 1991). An egocentric default has also been posited for dual-process models of perspective taking, which assume that both production and comprehension are initially egocentric, even when interlocutors are aware that successful communication requires taking the other's perspective into consideration (Epley, Keysar, van Boven, & Gilovich, 2004; Gilbert & Gill, 2000). Overcoming the egocentric bias to adjust to another perspective is a subsequent processing step that requires effort and practice (Keysar, Barr, & Horton, 1998). The difference that we observe between spoken and signed language modalities with respect to optimal viewpoint conventionalization (with spoken languages often adopting an addressee-viewpoint strategy and signed languages adopting an egocentric viewpoint strategy) suggests that the extent to which audience design can be observed in spatial perspective-taking tasks is dependent on the properties of the language modality in which the spatial descriptions are given. To better understand where and how audience design affects message planning requires measures that account for the relative cognitive burden of production and comprehension during communication, as well as the relative costs of unsuccessful communication (Keysar, Barr, Balin, & Brauner, 2000). The results of our study suggest that properties of the modality may modulate the distribution of these costs.

In sum, viewpoint alignment in the visual-spatial modality is necessary, but it does not come for free and is not automatic. Our results indicate a greater cost to accuracy when sign language producers relinquish their viewpoint, and the results support the hypothesis that viewpoint alignment must emerge with communicative experience over time. The nature of the visual-spatial modality likely constrains the pattern of viewpoint convergence, allowing signers, but not addressees, to remain egocentric.

ACKNOWLEDGEMENTS

We thank Alexandra del Solar, Jessica Gonzalez, Annemarie Kocab, Jenny Lu, Rachel Magid, Cindy O'Grady, Jennifer Petrich, Andrea Takahesu Tabori, Sarah Wylly, and Kate Zmich for research assistance. We also thank Ann Senghas for the stimuli used in Experiment 1. This research was supported by a postdoctoral fellowship funded by NIDCD Training Grant 5 T32 DC00041, a Wellesley College Provost's Office Grant, a fellowship from the Radcliffe Institute for Advanced Study (J.E.P), NIH/NIDCD grant R01 DC010997 (K.E.), and by a Marie Curie Fellowship MEIF-CT-2006-041886 (P.P.).

REFERENCES

Alibali MW, Hostetter AB. Mimicry and simulation in gesture comprehension. Behavioral and Brain Sciences. 2010; 33:433–434.

Amorim MA, Isableu B, Jarraya M. Embodied spatial transformations: "Body analogy" for the mental rotation of objects. Journal of Experimental Psychology. 2006; 13:527–547.

Brown, P.; Levinson, SC. Politeness: Some universals in language use. Cambridge University Press; Cambridge, UK: 1987.

Carlson SM, Mandell DJ, Williams L. Executive function and theory of mind: Stability and prediction from ages 2 to 3. Developmental Psychology. 2004; 40:1105–1122. [PubMed: 15535760]

Clark, HH. Space, time, semantics, and the child.. In: Moore, TE., editor. Cognitive development and acquisition of language. Academic Press; New York: 1973. p. 27-63.

Clark, HH. Arenas of language use. University of Chicago Press; Chicago: 1992.

Clark, HH.; Murphy, GL. Audience design in meaning and reference.. In: Leny, JF.; Kintsch, W., editors. Language and comprehension. North-Holland; Amsterdam: 1982. p. 287-299.

Dell, GS.; Brown, PM. Mechanisms for listener-adaptation in language production: Limiting the role of the "model of the listener.". In: Napoli, D.; Kegl, J., editors. Bridges between psychology and linguistics. Academic Press; New York: 1991. p. 105-129.

Emmorey, K. The confluence of space and language in signed languages.. In: Bloom, P.; Peterson, MA.; Nadel, L.; Garrett, MF., editors. Language and space. MIT Press; Cambridge, MA: 1996. p. 171-209.

Emmorey, K. Space on hand: The exploitation of signing space to illustrate abstract thought.. In: Gattis, M., editor. Spatial schemas and abstract thought. MIT Press; Cambridge, MA: 2001. p. 147-174.

Emmorey, K. The signer as an embodied mirror neuron system: Neural mechanisms underlying sign language and action.. In: Arbib, MA., editor. Action to language via the mirror neuron system. Cambridge University Press; Cambridge UK: 2006. p. 103-135.

Emmorey K, Klima E, Hickok G. Mental rotation within linguistic and non-linguistic domains in users of American Sign Language. Cognition. 1998; 68:221–246. [PubMed: 9852666]

Emmorey K, Tversky B. Spatial perspective choice in ASL. Sign Language and Linguistics. 2002; 5:3–25.

Epley N, Keysar B, van Boven L, Gilovich T. Perspective taking as egocentric anchoring and adjustment. Journal of Personality and Social Psychology. 2004; 87:327–339. [PubMed: 15382983]

Galati A, Avraamides MN. Flexible spatial perspective-taking: conversational partners weigh multiple cues in collaborative tasks. Frontiers in Human Neuroscience. 2013; 7(618):1–16. doi: 10.3389/fnhum.2013.00618. [PubMed: 23355817]

Galati A, Michael C, Mello C, Greenauer NM, Avraamides MN. The conversational partner's perspective affects spatial memory and descriptions. Journal of Memory and Language. 2013; 68:140–159.

Gilbert DT, Gill MJ. The momentary realist. Psychological Science. 2000; 11:394–398. [PubMed: 11228910]

Goldin-Meadow S, McNeill D, Singleton J. Silence is liberating: Removing the handcuffs on grammatical expression in the manual modality. Psychological Review. 1996; 103:34–55. [PubMed: 8650298]

Herrmann T, Bürkle B, Nirmaier H. Zur hörerbezogenen Raumreferenz: Hörerposition und Lokalisationsaufwand. Sprache & Kognition. 1987; 6(3):126–137.

Horton WS, Keysar B. When do speakers take into account common ground? Cognition. 1996; 59:91–117. [PubMed: 8857472]

Kessler K, Thomson LA. The embodied nature of spatial perspective taking: embodied transformation versus sensorimotor interference. Cognition. 2010; 114:72–88. [PubMed: 19782971]

Keysar B, Barr D, Horton WS. The egocentric basis of language use: Insights from a processing approach. Current Directions in Psychological Science. 1998; 7:46–50.

Keysar B, Barr D, Balin J, Brauner J. Taking perspective in conversation: The role of mutual knowledge in comprehension. Psychological Science. 2000; 11:32–38. [PubMed: 11228840]

Keysar B, Lin S, Barr DJ. Limits on theory of mind use in adults. Cognition. 2003; 89:25–41. [PubMed: 12893123]

Levelt, WJM. Speaking: From intention to articulation. The MIT Press; Cambridge, MA: 1989.

Levinson, SC. Space in language and cognition: explorations in cognitive diversity. Cambridge University Press; Cambridge UK: 2003.

Lu C-H, Proctor RW. The influence of irrelevant location information on performance: A review of the Simon and spatial Stroop effects. Psychonomic Bulletin & Review. 1995; 2:174–207. [PubMed: 24203654]

MacLeod CM. Half a century of research on the Stroop effect: An integrative review. Psychological Bulletin. 1991; 109:163–203. [PubMed: 2034749]

Mainwaring SD, Tversky B, Ohgishi M, Schiano DJ. Descriptions of simple spatial scenes in English and Japanese. Spatial Cognition and Computation. 2003; 3:3–42.

Miller, GA.; Johnson-Laird, PN. Language and perception. Harvard University Press; Cambridge, MA: 1976.

MPI for Psycholinguistics Field Manual. Space stimuli kit 1.2. Levinson, Stephen C., editor. Max Planck Institute for Psycholinguistics; Nijmegen: Nov. 1992 p. 511992

Perniss PM. Achieving spatial coherence in German Sign Language narratives: The use of classifiers and perspective. Lingua. 2007; 117:1315–1338.

Perniss P, Özyürek A. Channels of expression differentially modulate iconicity and embodiment in co-speech gesture and sign language. under revision.

Piaget, J.; Inhelder, B. The child's conception of space. Routledge & Kegan Paul; London: 1956.

Pickering M, Garrod S. An integrated theory of language production and comprehension. Journal of Behavior and Brain Sciences. 2013; 36:375–392.

Pyers, J.; Perniss, P.; Emmorey, K. Viewpoint in the visual-spatial modality. Paper presented at the 30th annual convention of the German Society of Linguistics workshop on Gestures: A comparison of signed and spoken languages. Bamberg; Germany: Feb. 2008

Pyers J, Senghas A. Language promotes false-belief understanding: Evidence from learners of a new sign language. Psychological Science. 2009; 20:805–812. [PubMed: 19515119]

Pyers J, Shusterman A, Senghas A, Spelke E, Emmorey K. Evidence from an emerging sign language reveals that language supports spatial cognition. Proceedings of the National Academy of Science. 2010; 107:12116–12120.

Quinto-Pozos D, Singleton JL, Hauser PC, Levine SC, Garberoglio CL, Hou L. Atypical signed language development: A case study of challenges with visual-spatial processing. Cognitive Neuropsychology. 2013; 30:332–359. [PubMed: 24344817]

Schober MF. Spatial perspective-taking in conversation. Cognition. 1993; 47:1–24. [PubMed: 8482069]

Schober MF. Speakers, addressees, and frames of reference: Whose effort is minimized in conversations about location? Discourse Process. 1995; 20:219–247.

Senghas, A. Delivered at the Child Language Seminar. City University; London: Jun 25. 2010a From gestures to grammar – How learners created Nicaraguan Sign Language. Keynote address.. 2010

Senghas A. The emergence of two functions for spatial devices in Nicaraguan Sign Language. Human Development. 2010b; 53:285–300.

Stroop JR. Studies of interference in serial verbal reactions. Journal of Experimental Psychology. 1935; 18:643–662.

Talmy, L. The representation of spatial structure in signed and spoken language.. In: Emmorey, K., editor. Perspectives on classifier constructions. Lawrence Erlbaum Associates; Mahwah, New Jersey: 2003. p. 169-195.

Tversky B, Hard B. Embodied and disembodied cognition: Spatial perspective-taking. Cognition. 2009; 110:124–129. [PubMed: 19056081]

Wilson M. Perceiving imitable stimuli: Consequences of isomorphism between input and output. Psychological Bulletin. 2001; 127:543–553. [PubMed: 11439711]

Wraga M, Creem SH, Proffitt DR. The influence of spatial reference frames on imagined object- and viewer rotations. Acta Psychologica. 1999; 102:247–264. [PubMed: 10504883]

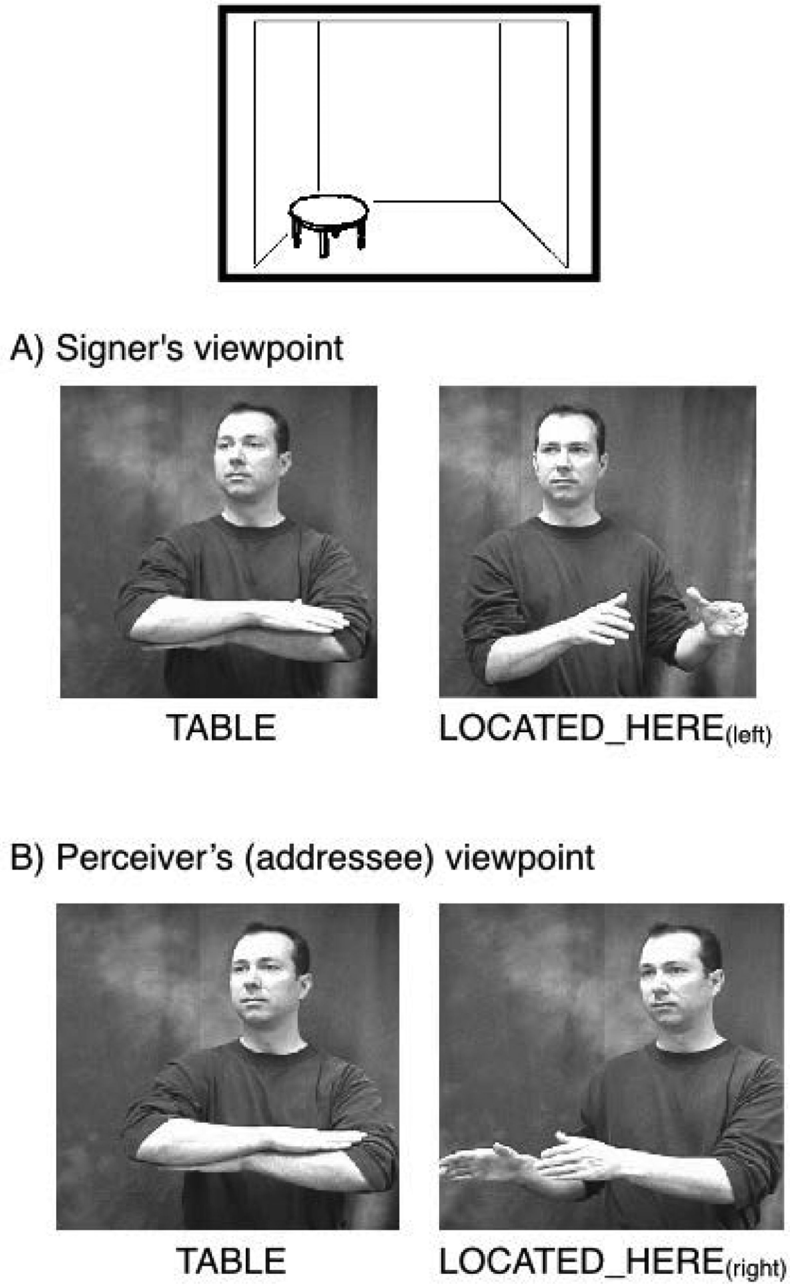

Figure 1.

ASL description of the location of a table in the picture a) from the signer's egocentric perspective as he views the picture and b) from the addressee's viewpoint as he/she views the signs.

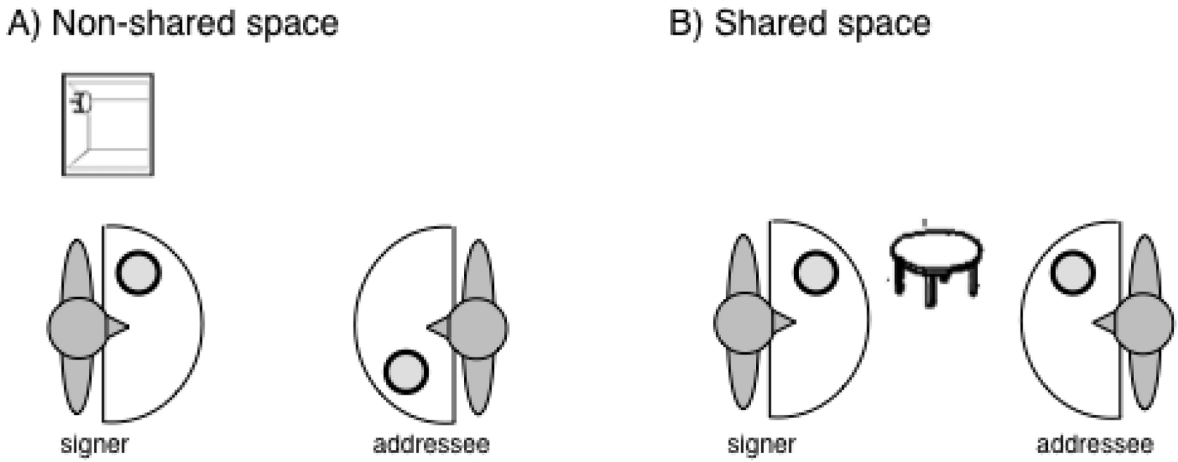

Figure 2.
(A) When the signer talks about non-present referents or non-jointly viewed referents, the signer describes the spatial array from his own viewpoint and the addressee interprets the utterance non-egocentrically. (B) When signer and addressee jointly view an environment their different vantage points on the scene map to the same, shared physical space. The description of the scene in sign space by the signer or the addressee matches the real-world locations of the objects.

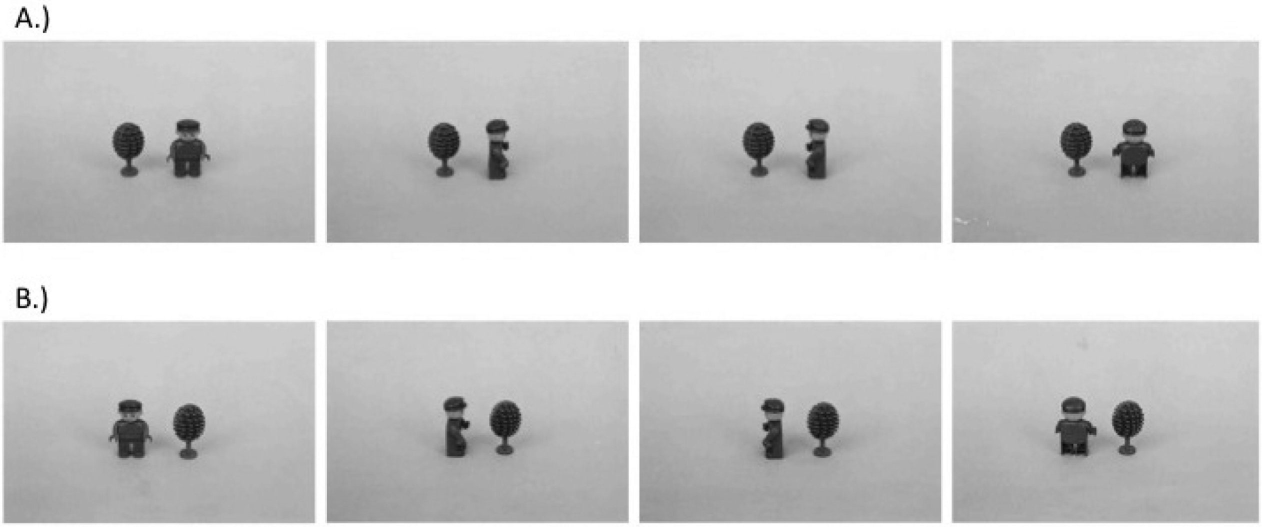

Figure 3.
The eight test pictures used in Experiment 1. The orientation of the man was not coded. Therefore, the four pictures in Set A were considered correct matches for each other, and the four pictures in Set B were considered correct matches for each other.

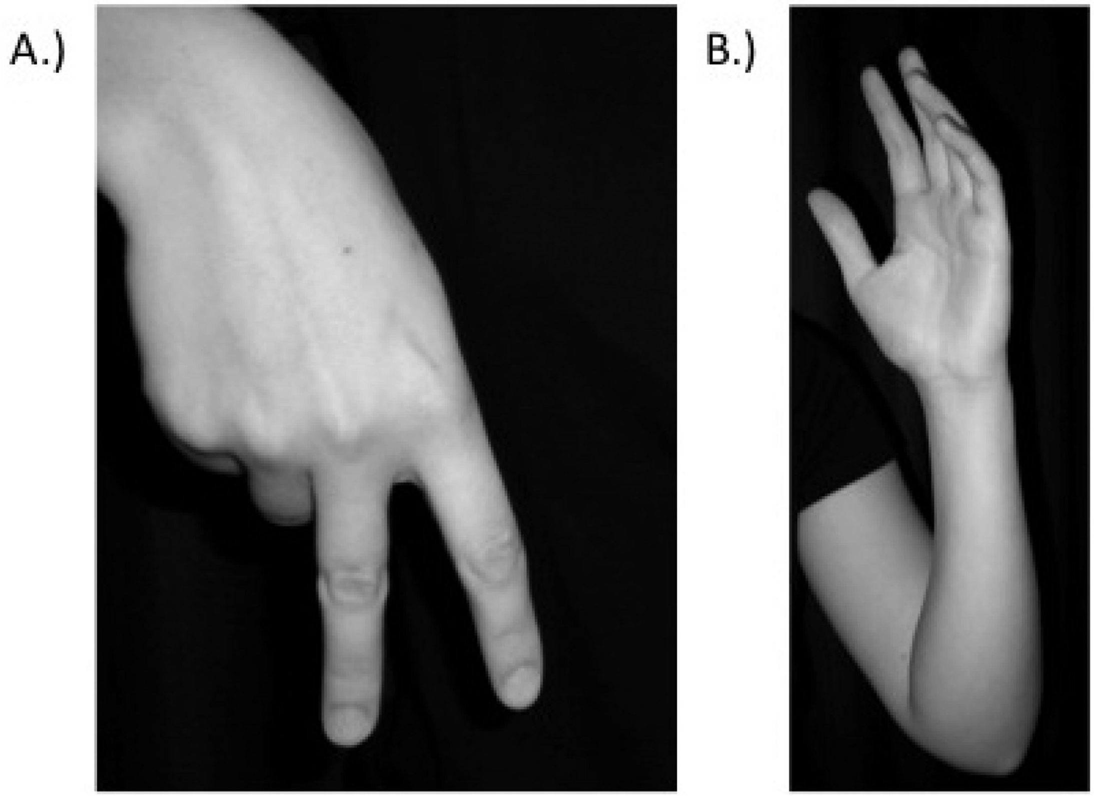

Figure 4.
The two gestures taught to hearing sign-naïve participants in Experiment 1 to represent A) the man and B) the tree.

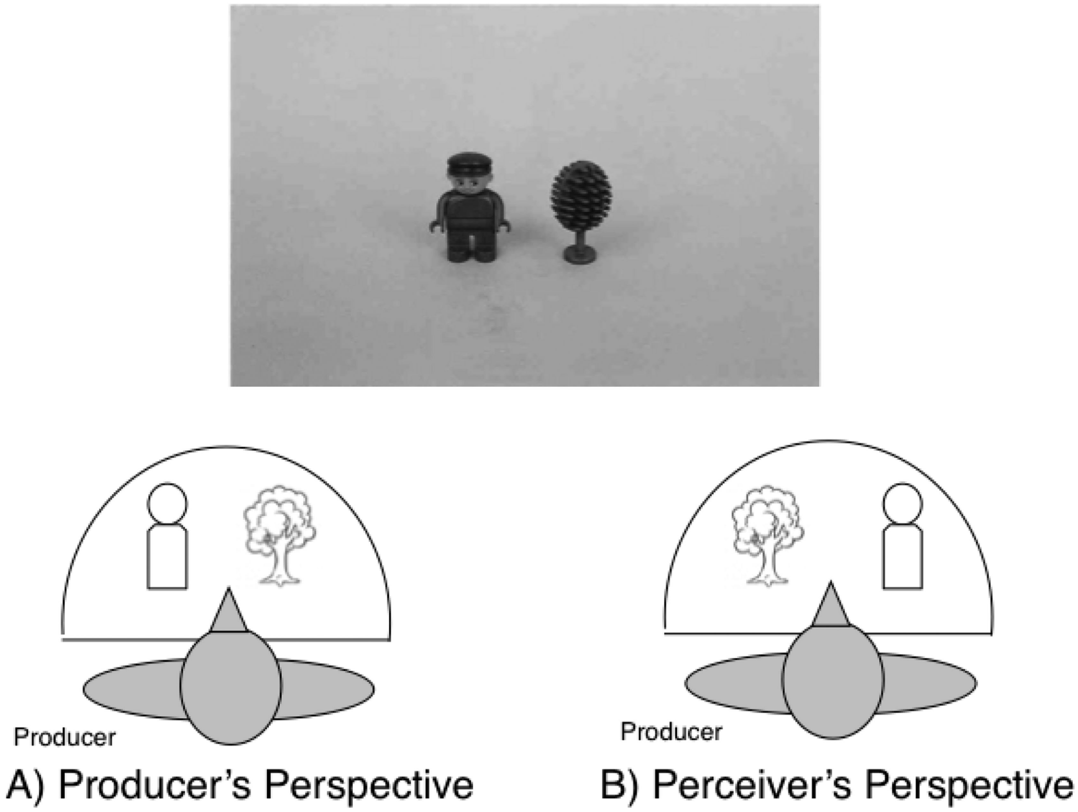

Figure 5.
Diagram of how the Producer's utterances were coded for viewpoint. A) Signers or gesturers were coded as using the “Perceiver's Perspective” when they mapped the relative locations of the sign or gesture for ‘man’ and the sign or gesture for ‘tree’ directly onto the locations represented in the stimulus. In this case, the man is to the left of the tree from the Producer's perspective, and the Producer locates the man on his left side and the tree on his right side. B) Signers or gesturers were coded as using the “Producer's Perspective” when they reversed the relative locations of the sign or gesture for ‘man’ and the sign or gesture for ‘tree’ such that when the Perceivers viewed the manual production, the manual production mapped directly onto the spatial scene viewed by the Perceiver. In this case the man is to the left of the tree from the Producer's perspective, and the Producer locates the man on his right side (the Perceiver's left side) and the tree on his left side (the Perceiver's right side).

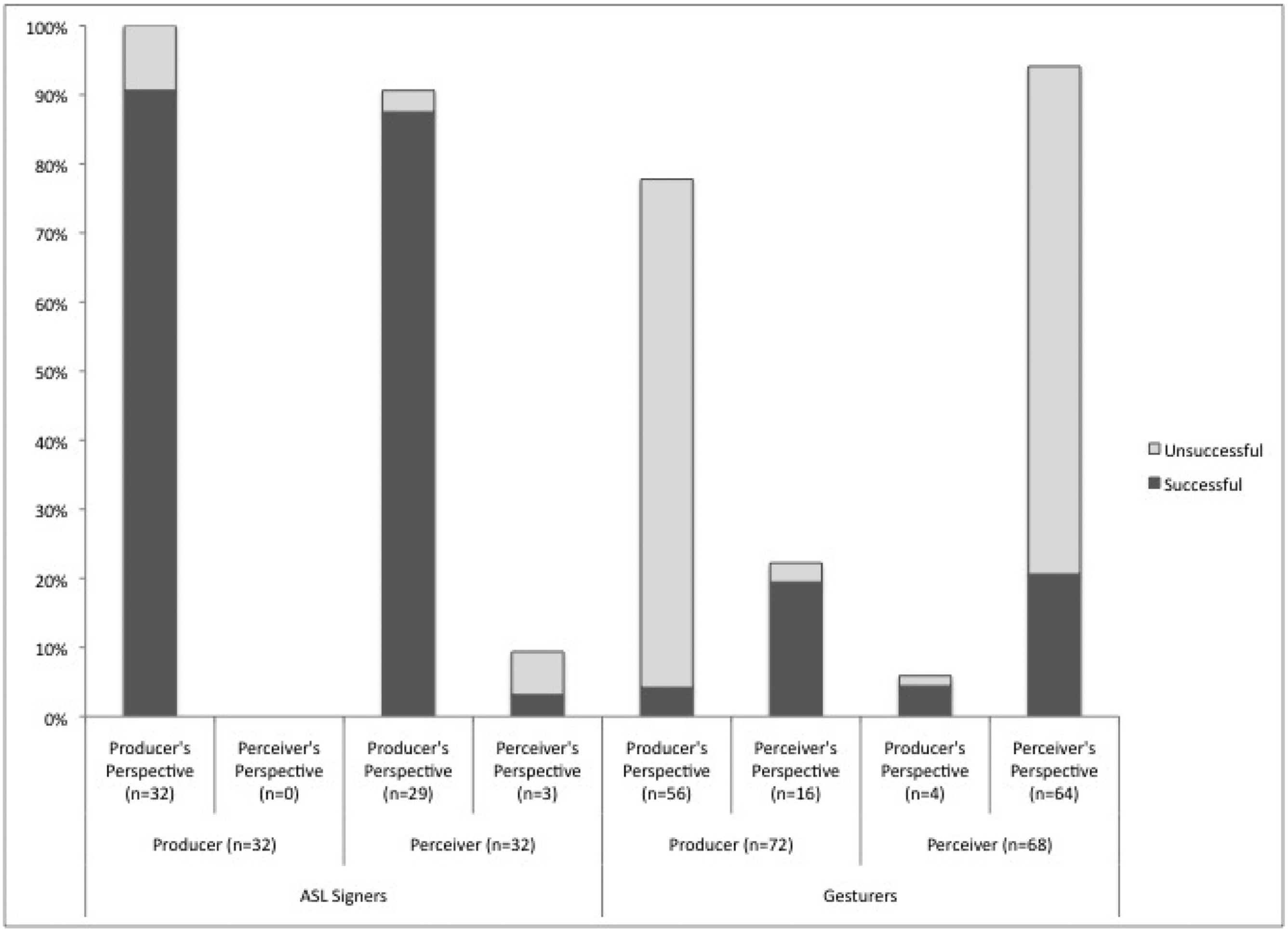

Figure 6.

Percentage of successful and unsuccessful picture selections for each viewpoint strategy presented by Producers and interpreted by Perceivers in Experiment 1.

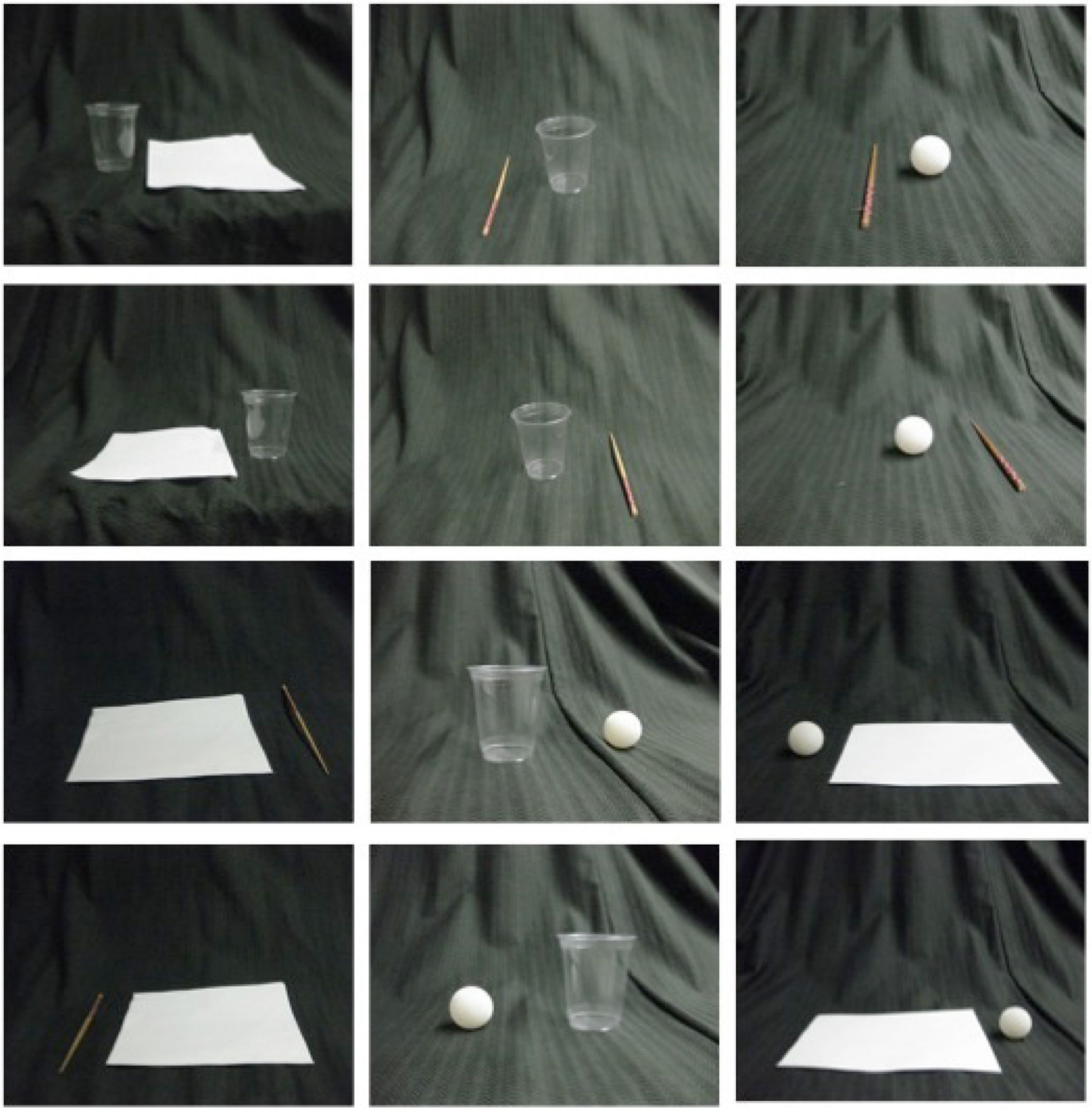

Figure 7.
The Perceiver's layout of the twelve test pictures used in Experiment 2.

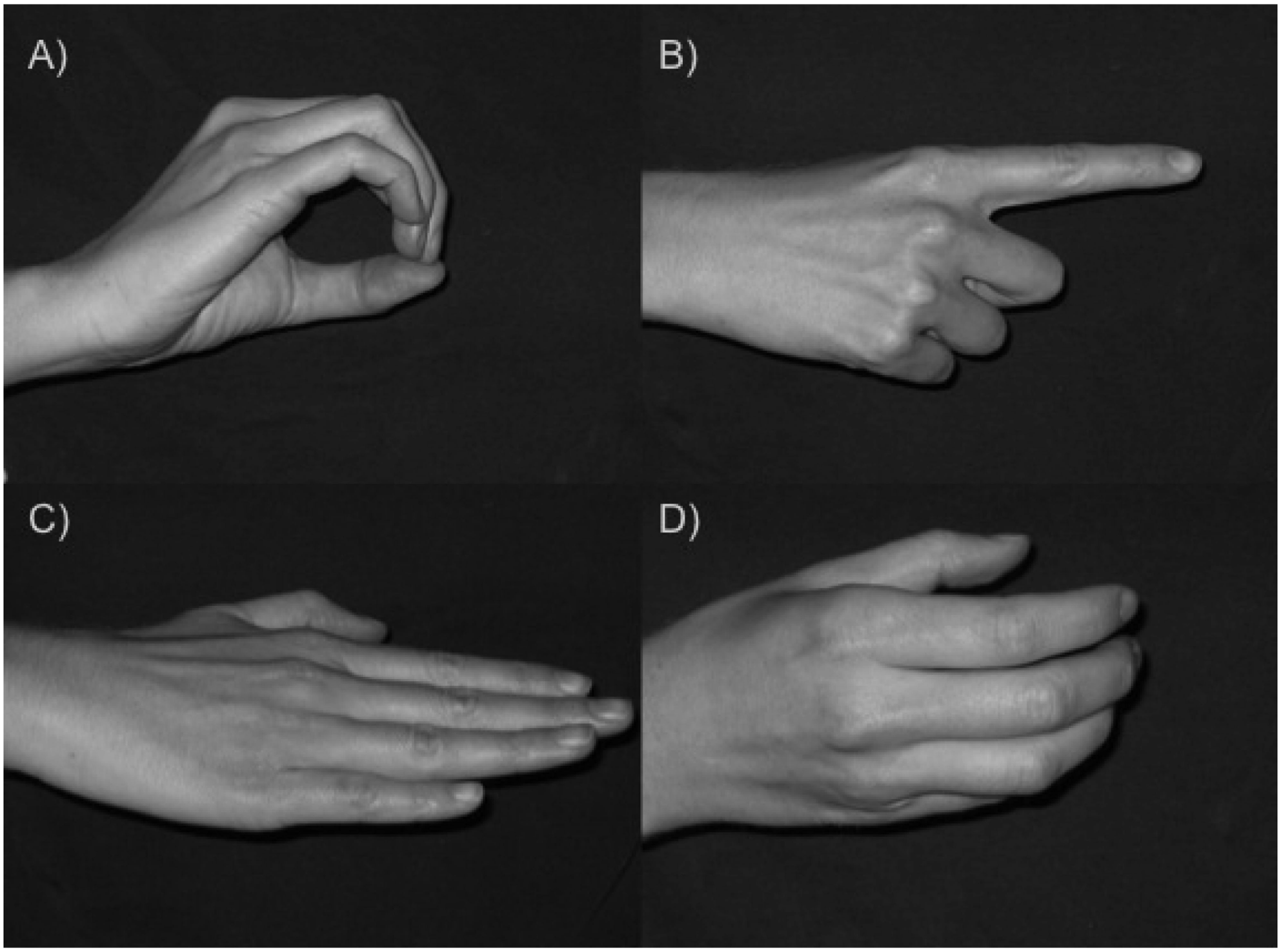

Figure 8.
The four gestures taught to hearing sign-naïve participants in Experiment 2 to represent A) the ball, B) the chopstick, C) the paper, and D) the cup.

Table 1

Means (Standard Deviations) of Perspective Choices in Experiment 1 and in the Unassigned Condition of Experiment 2

		Producer		Perceiver		
Group	**Pairs (*N*)**	**Producer's Perspective (%)**	**Perceiver's Perspective (%)**	**Producer's Perspective (%)**	**Perceiver's Perspective (%)**	**Correct Picture Selection (#)**[a]
Experiment 1						
ASL Signers	4	100.0 (0.00)	0.0 (0.00)	90.6 (6.25)	9.4 (6.25)	7.3 (0.50)
Gesturers	9	77.8 (34.67)	22.2 (34.67)	5.6 (6.59)	88.9 (14.58)	1.9 (2.67)
Experiment 2						
Gesturers Unassigned Perspective	11	82.2 (28.41)	17.8 (28.41)	35.6 (43.32)	64.4 (43.32)	5.2 (4.77)

[a]Out of 8 possible for Experiment 1 and 12 possible for Experiment 2.

Table 2

Means (Standard Deviations) and Inferential Statistics of Response Variables in the Experimental Conditions of Experiment 2

		Producer			Perceiver		
Condition	**Pairs (*N*)**	**Correct Viewpoint Presented (%)**	**First trial correct production time (s)**	**Average Correct Production time (s)**	**Correct Viewpoint Interpreted (%)**	**First trial correct selection time (s)**	**Average Correct Selection Time (s)**
Producer's Perspective	17	98.6 (3.12)	2.8 (1.49)	3.1 (1.55)	80.4 (36.08)	3.6 (1.16)	3.0 (.75)
Perceiver's Perspective	19	78.5 (34.74)	4.2 (2.26)	3.8 (1.72)	85.53 (31.3)	2.90 (1.8)	2.73 (.2)
Producer's vs Perceiver's Conditions	*t*	2.51[a]	2.04	0.46	0.46	1.07	1.23
	df	18.32	26.00	34.00	34.00	22.00	32.00
	p	.02	.05	.65	.65	.30	.23
	d	.81	.75	.39	.15	.27	.55

[a]Assuming unequal sample variances, $F(18,16) = 125.67$, $p < .0001$