Empirical Article

Eye Gaze During Comprehension of American Sign Language by Native and Beginning Signers

Karen Emmorey
San Diego State University

Robin Thompson
University College London

Rachael Colvin
University of California, San Diego

An eye-tracking experiment investigated where deaf native signers (N = 9) and hearing beginning signers (N = 10) look while comprehending a short narrative and a spatial description in American Sign Language produced live by a fluent signer. Both groups fixated primarily on the signer’s face (more than 80% of the time) but differed with respect to fixation location. Beginning signers fixated on or near the signer’s mouth, perhaps to better perceive English mouthing, whereas native signers tended to fixate on or near the eyes. Beginning signers shifted gaze away from the signer’s face more frequently than native signers, but the pattern of gaze shifts was similar for both groups. When a shift in gaze occurred, the sign narrator was almost always looking at his or her hands and was most often producing a classifier construction. We conclude that joint visual attention and attention to mouthing (for beginning signers), rather than linguistic complexity or processing load, affect gaze fixation patterns during sign language comprehension.

Because sign language is perceived visually, the eye movements and gaze position of an addressee allow us to make inferences about the uptake of linguistic information in real time. We used eye-tracking technology to determine whether eye gaze movement during sign language comprehension is affected by information content, as has been found for eye movements during reading and in “visual world” experiments. For example, when viewing a visual scene, eye movements are closely time locked to object information presented in a spoken utterance (e.g., Tanenhaus, Spivey-Knowlton, Eberhard, & Sedivy, 1995). During silent reading, gaze fixation patterns are frequently used as a measure of local processing difficulty for words or phrases. For example, increased fixation times and regressive (backtracking) eye movements typically indicate that a reader is having difficulty with a particular region of text (for review, see Staub & Rayner, 2007). We therefore hypothesized that patterns of gaze fixation and movement might provide a measure of processing difficulty for sign language comprehension. Just as readers fixate longer and backtrack over regions of difficult text, it is possible that sign perceivers shift fixation toward the hands when comprehending complex linguistic structures that are conveyed by the manual articulators. Examinations of videotaped signed interactions, as well as introspective data from native signers, suggest that addressees maintain a relatively steady gaze toward the person signing (Baker & Padden, 1978; Siple, 1978). However, there is very little evidence regarding precisely where addressees look when processing sign language and whether there are specific changes of fixation with respect to the signer’s face and hands at particular points during language comprehension.

We would like to thank Hlsa Borinstein, Melissa Herzig, Franco Korpics, Taylor Mayer, and Steven McCallough for help running the study and Jeane Kim, Heather Larrabee, Renee McMorris, Rachel Morgan, and Monica SolioniCo for help with data analysis. No conflicts of interest were reported. Correspondence should be sent to Karen Emmorey, Laboratory for Language and Cognitive Neuroscience, 6495 Alvarado Road, Suite 200, San Diego, CA 92102 (e-mail: kemmorey@mail.sdsu.edu).
Previously, Muir and Richardson (2005) used eye tracking to explore the gaze patterns of deaf users of British Sign Language (BSL) while they watched videotaped signing. Deaf signers viewed three short video clips of signed stories that were selected to include a wide range of fine and gross motor movements. Participants in this study fixated on the signer’s face between 61% and 99% of the time across the three video clips. Examination of hand movements in the videos suggested that the following factors caused shifts in gaze away from the face and toward the signer’s hands or body: (a) signs close to the face (gaze is drawn to the hands), (b) “expansive” signs in the lower body region, and (c) movement of the signer within the video scene. In one of the video clips, participants tended to fixate on the upper body of the signer, rather than on the face, and Muir and Richardson (2005) hypothesized that the wider and more rapid movements produced by the signer may have caused gaze to fall on the upper body to permit a range of movements to be processed, while keeping the lower part of the face in foveal (high-resolution) vision.

However, watching a signer on videotape may be quite different from watching a live signer as an addressee. Video is two dimensional, and because the relative size of the signer is smaller than in real life, a more central fixation point may be a better strategy when viewing sign language within a smaller field. In addition, eye contact may be much more critical when watching a narrative produced by a signer who is actually present. Maintaining eye contact is a signal to the signer that he or she has the floor, whereas a shift in gaze away from the signer can indicate a desire for a conversational turn (Baker, 1977). We hypothesized that for live interactions, addressees would be less likely to fixate on the upper body of the signer and more likely to fixate on the signer’s eyes.

In addition, we examined whether and how eye gaze behaviors of beginning hearing signers differ from those of deaf native signers. We hypothesized that beginning signers would be more likely to shift their gaze to the hands of the signer and that they might exhibit a smaller “perceptual span.” In reading, perceptual span refers to that portion of the text from which useful information is obtained during a single fixation and is affected by the nature of a language’s orthography (Rayner & Sereno, 1994). Perceptual span is larger to the right of fixation for English readers (English is read left to right), to the left of fixation for Hebrew (read right to left), and is smaller for Chinese (read vertically). We adopt this notion of perceptual span to refer to the distance between fixation and where the signed signal appears within the visual field. If perceptual span is affected by experience as it is for reading, we hypothesize that beginning signers might have a narrower perceptual span compared to native signers, and thus they may need to move fixation off the eyes in order to keep the signer’s hands within this smaller perceptual span. In addition, deafness may affect perceptual span. Several studies have found that auditory deprivation leads to an enhanced ability to detect and attend to motion in the periphery of vision (Bavelier et al., 2000; Bavelier, 2002). Thus, we predict that beginning signers who are hearing will spend less time fixating on the face, will fixate on the lower part of the face or on the upper body of the signer, and will also shift fixation to the hands more frequently than native deaf signers.

In this study, a fluent American Sign Language (ASL) signer first told a simple story and then described the layout of a town in ASL to either a deaf native signer or to a hearing beginning signer. The two narratives differed in the amount of fingerspelling and in the number of locative classifier constructions. We predicted that these factors would affect where the addressee looked during sign comprehension. Although Muir and Richardson (2005) reported that participants’ gaze did not move to the hands during BSL fingerspelling, ASL uses a one-handed fingerspelling system that is produced more laterally, with the dominant hand displaced toward the side of the body. Thus, ASL fingerspelled words that are produced near the periphery of vision might be more likely to draw gaze away from the face than the more centrally positioned BSL fingerspelled words.

We also predicted that addressees would be more likely to shift their gaze toward the signer’s hands for locative classifier constructions because the location of the hand(s) in space conveys critical semantic information and because signers often look at their own hands when producing locative classifier constructions. Gullberg and Holmqvist (1999, 2006) found that during
spoken conversations, addresses tended to shift gaze from the face to "autofixed" gestures, that is, gestures for which the speaker looked at his or her own hand(s). Speakers may gaze toward their own hand(s) in order to call attention to their gestures. Gullberg and Holmqvist (2006) reported that it was common for speakers who fixated on their own gesture to look back up at their addressee to check whether he or she was also looking at their gesture. We hypothesized that sign addressees might also be sensitive to a signer’s autofixed gaze as a way to draw attention to a particular sign and that sign perception might pattern like gesture perception in this respect. Thus, we examined whether shifts of gaze away from the face coincided with autofixed signs.

In sum, we investigated where native and beginning signers look while comprehending an ASL narrative and a spatial description, examining the location of fixation on the face (e.g., on or near the eyes, mouth, or forehead), the frequency of gaze shifts away from the face, and the content of signing when gaze shifted away from the face.

**Methods**

**Participants**

Nine deaf native ASL signers participated (four males, five females; mean age = 25.9 years; SD = 2.6). All were born into Deaf signing families and used ASL as their preferred and primary language. Ten hearing beginning signers participated (two males, eight females; mean age = 20.4 years; SD = 1.26). At the time of testing, beginning signers had completed between 9 and 15 months of ASL instruction (three to five ASL courses) at the University of California, San Diego, with 6 hr of in-class time per week. Following the experiment, nine of the beginning signers and seven of the native signers were given a general test of ASL comprehension ability in which they watched two ASL stories followed by detailed multipart questions. The beginning signers had a mean score of 16.22 (SD = 12.39) out of a possible 72, and the mean score for the native signers was 52.86 (SD = 8.65).

**Procedure**

While wearing a head-mounted eye tracker, participants watched one of five native signers producing two narratives. The first narrative was a simple story about children painting each other while their mother’s back was turned (the Paint Story, which has been used in studies of sign language acquisition; e.g., Bellugi, van Hock, Lillo-Martin, & O’Grady, 1993). The second signed narrative was a description of the spatial layout of a town (from the Town map, which has been used in studies of spatial language; Emmorey, Tversky, & Taylor, 2000). The mean length of the Paint Story narrative was 1.35 min (SD = 0.06) when told to novice signers and 0.98 min (SD = 0.46) when told to native signers, and this difference in story length was not significant, t(17) = 1.38, p = .191. After the narrative, participants were asked five questions to determine whether they comprehended the story. Beginning signers generally understood the story, answering a mean of 4.44 questions correctly (SD = 0.73). Native signers answered a mean of 4.78 questions correctly (SD = 0.44). The mean length of the Town description was 3.28 min (SD = 1.46) when told to novice signers and 2.42 min (SD = 0.96) when told to native signers, and description length did not differ significantly between groups, t(17) = 1.73, p = .106. The trend for slightly longer narratives for the beginning signers may reflect a tendency to sign more slowly for a beginning compared to native signing addressee. After the Town layout was described, participants were asked to draw a map of the landmarks in the Town. Each correctly placed landmark was scored, and the two groups did not differ significantly in their understanding of the Town layout, t(17) = 0.37, ns. The beginning signers remembered and located 74.85% of the landmarks correctly (SD = 13.42%), and the native signers had a mean of 76.9% correct (SD = 17.27%).

Participants’ eye movements were monitored using the iView system from SensoMotoric Instruments, Inc. The iView system provides video-based evaluation of gaze position using real-time image processing. The SMI Head-mounted Eye-tracking Device (HED) consists of two miniature video cameras, an infrared illuminator, and a double-sided dichroic mirror, all attached to a lightweight bicycle helmet. The HED weighs 450 g is adjustable to all head sizes and has a resolution of 0.1 degrees. The dichroic mirror reflects infrared light, but it appears transparent to
the participant. The scene camera shows the participant’s field of view as reflected on the outside of the mirror, and the eye camera tracks eye movements from the reflected image of the eye on the inside of the mirror. In the resulting video, a cursor indicating the participant’s eye position is superimposed onto the image of the participant’s field of view. A major advantage of head-mounted eye tracking is that participants’ head movements are not restricted by a chin rest, allowing for much more natural interaction.

For each participant, we calculated the percentage of time that they fixated on or near the face of the signer. For the time that fixation was on the face, we calculated the mean percentage of time that they fixated on the following features: the “upper face” (on the upper forehead or on a location just above the upper part of the face), “the eyes” (including the lower forehead, between the two eyes, and the upper part of the nose), “the mouth” (including the tip of and just below the nose and the upper chin), and the “lower face” (the edge of the chin or on a location just below the lower part of the face). Finally, when a participant’s gaze shifted away from the face toward the signer’s hands or body, we coded whether the signer was producing (a) a lexical sign, (b) a fingerspelled word, or (c) a classifier sign and whether the signer was autifixating on the sign.

Results

The percentage of time that participants fixated on the face was entered into a two-way analysis of variance (ANOVA), with participant group (native signer, beginning signer) as a between-group factor and narrative type (Paint Story, Town Description) as a within-group factor. Consistent with previous videotape data, the eye-tracking results revealed that deaf signs looked primarily at the signer’s face. Surprisingly, however, the beginning and native signs did not differ from each other with respect to the amount of time that gaze remained on the face, 87.9% and 88.8%, respectively, $F(1, 17) < 1$ (see Table 1). In addition, the results revealed a main effect of narrative type, $F(1, 17) = 12.67, p = .002$, that did not interact with participant group, $F(1, 17) < 1$. Both native and beginning signs spent more time looking off the face during the Town description (18.56% of the time) than during the Paint Story narrative (only 4.67% of the time).

The percentage of time that participants gazed at different locations on the face was entered into a repeated-measures ANOVA, with participant type (native, beginning signer) as a between-subject variable and narrative type (Paint Story, Town Description) as a within subjects variable. These data are presented in Table 2. There was a significant main effect of gaze location, $F(1, 3) = 13.126, p < .001$, with both participant groups fixating more often on or near the eyes (37.28%) and the mouth (41.72%) than just above the forehead (2.63%) or just below the face (18.4%). Narrative type did not interact with gaze location or with participant group. However, there was a significant interaction between gaze location and participant group, $F(1, 3) = 2.82, p < .05$. As illustrated in Figure 1, beginning and native signs differed with respect to where they looked on the face while comprehending ASL. Overall, beginning signs spent significantly more time looking at the mouth area than native signers, 53.14% versus 30.31%, $F(1, 17) = 8.28, p < .02$; whereas, native signers fixated more on or near the signer’s eyes than beginning signers, 45.64% versus 28.92%, although this difference was not significant, $F(1, 17) = 1.84, p = .190$.

For each narrative, we coded the number of shifts of gaze away from the face and whether the signer was producing a fingerspelled word, a classifier construction, or a lexical sign when the participant’s gaze moved off the face. Two native signers never shifted their gaze from the face and therefore were not included in the analysis. The results are shown in Table 3. Novice signers shifted their gaze away from the face significantly more often than native signers, $F(1, 15) = 5.62, p < .05$, and the group difference in number of gaze shifts was greatest for the Town Description, $F(1, 15) = 4.62, p < .05$, for the interaction between participant group and narrative type.

<table>
<thead>
<tr>
<th>Table 1</th>
<th>Mean percentage of time that participants fixated on the signer’s face for each narrative</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Paint Story</td>
</tr>
<tr>
<td>Native signs</td>
<td>94.97 (8.55)</td>
</tr>
<tr>
<td>Beginning signs</td>
<td>95.67 (5.40)</td>
</tr>
</tbody>
</table>

Note: Standard deviations are in parentheses.
As expected, more gaze shifts occurred during the Town Description than during the Paint Story, $F(1, 15) = 14.37, p = .002$, and this pattern is largely due to the fact that the Town Description contained more classifier constructions. Both groups were significantly more likely to shift their gaze toward classifier constructions than toward lexical signs or fingerspelled words (see Table 3), $F(2, 15) = 30.20, p < .001$. In contrast to our predictions, however, sign perceivers did not routinely shift their gaze toward the signer’s hand to perceive fingerspelled words.

Finally, we coded whether the signer was looking toward his or her hands when participants shifted their gaze away from the face. The results indicated that the signer was almost always looking at the sign(s) that he or she was producing when participants shifted their gaze from the face. Of the total of 211 gaze shifts (157 by beginning signers and 54 by native signers), only three gaze shifts occurred when the signer was looking at the participant, rather than at his or her own hands (all three of these gaze shifts were made by the same participant, a native signer). Thus, nearly 99% of gaze shifts occurred when the signer was producing an autofixated sign.

### Discussion

Replicating previous eye-tracking studies (Agrafiotis, Canagarajah, & Bull, 2003; Muir & Richardson, 2005), deaf signers fixated almost exclusively on the face during comprehension of signed narratives. Contrary to our predictions, beginning signers did not spend less time fixating on the face than native signers. We propose that several factors conspire to create this pattern of eye gaze for the perception of sign language. Most importantly, sign comprehension requires visual perception of unpredictable, dynamic information. By fixating on the face, sign perceivers need not predict where the hand(s) will move next and can take in dynamic information from the periphery of vision. In addition, grammatical information is conveyed on the face, and subtle changes in facial expression would be harder to perceive in the periphery if the eyes were to move toward signing space. Finally, in American culture, it is accepted (and often expected) for an addressee to gaze at the speaker’s face. Together, these factors create a strong predisposition to fixate on the signer’s face such that beginning signers with only months of signing experience learn very quickly to fixate on the face during sign comprehension.

However, the patterns of gaze fixation on the face differed significantly between beginning and native signers. Beginning signers tended to fixate on or near the signer’s mouth, whereas native signers tended to fixate on or near the eyes. We hypothesize that beginning signers looked more at the mouth in order to pick up additional information conveyed by mouthing. During both narratives, signers frequently produced mouth patterns that look like English words (or a part of an English word) while simultaneously producing the corresponding manual sign. Information about the English translations of ASL signs would be particularly useful to beginning signers, and it is likely that the sign storytellers mouthed more for the beginning

### Table 2

Mean percentage of time that participants fixated on locations that were on or near the signer’s face for each narrative (Paint Story and Town Description)

<table>
<thead>
<tr>
<th></th>
<th>Just above the face</th>
<th>On/near the eyes</th>
<th>On/near the mouth</th>
<th>Just below the face</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Paint</td>
<td>Town</td>
<td>Paint</td>
<td>Town</td>
</tr>
<tr>
<td>Native signer</td>
<td>2.30 (1.06)</td>
<td>3.13 (1.60)</td>
<td>51.84 (10.0)</td>
<td>39.43 (9.34)</td>
</tr>
<tr>
<td>Beginning signer</td>
<td>2.41 (1.01)</td>
<td>2.69 (1.52)</td>
<td>33.35 (9.48)</td>
<td>24.49 (8.86)</td>
</tr>
</tbody>
</table>

*Note. Standard deviations are in parentheses.*

### Figure 1

Mean percent of gaze toward locations on the signer’s face for each participant group. Bars indicate standard error.
signers than for the native signers. Frequent mouthing by a sign narrator may cause an addressee to shift gaze toward the mouth, and it is possible that native signers might also gaze more frequently at the mouth if the sign narrator produced a substantial amount of English mouthing (as might be expected for orally trained, deaf late learners of ASL).

As predicted, native signers tended to gaze on or near the eyes of the signer (see Figure 1). Thus, during live interactions (as opposed to watching a videotape), native signers show a preference for eye contact. Maintaining steady gaze toward the signer’s eyes expresses social cues and may also enhance language comprehension. Socially, eye contact signals attention by the addressee and indicates that the sign narrator has the floor because gaze away from the signer would signal the desire for a conversational turn (Baker, 1977). With respect to language comprehension, gaze toward the signer’s eyes provides a stable fixation point for sign perception and enables the addressee to easily detect changes in eye gaze that convey grammatical and referential information. For example, changes in eye gaze signal role shifts within a narrative (Bahan & Supalla, 1995), and gaze direction can be used for deictic reference (Baker & Padden, 1978). Eye gaze also functions to mark verb agreement in ASL (Neidle, MacLaughlin, Bahan, & Lee, 2000; Thompson, Emmorey, & Kluender, 2006), and changes in eyebrow configuration mark various syntactic structures (e.g., conditional clauses, WH clauses, topics, and yes–no questions; Liddell, 1980).

We did not find strong evidence for a narrower perceptual span for beginning signers compared to native signers. Although beginning signers fixated more on the mouth area (including the upper chin), they did not fixate more below the face (see Figure 1) and also did not fixate on the upper body. Beginning signers did shift their gaze toward the hands more frequently than native signers, but their gaze did not remain on or near the hands for a significant amount of time, evidenced by the fact that beginning signers did not spend significantly more time gazing off the face than native signers. Finally, the fact that gaze shifts were relatively infrequent across a narrative (see Table 3) suggests that beginning signers did not shift their gaze to the hands to compensate for a narrower perceptual span.

Finally, the most striking finding of this study was that nearly all gaze shifts occurred when the sign narrator was looking at his or her own hands. Gaze patterns for sign language perception thus paralleled those for the perception of co-speech gesture. Gullberg and Holmqvist (2006) found that addressees primarily gazed at the speaker’s face and rarely shifted gaze to the speaker’s hands—the vast majority of co-speech gestures (93%) did not attract fixation. However, when gaze did shift to the speaker’s gesture, the speaker was frequently also looking at his or her hands while producing the gesture. It is likely that signers and speakers look toward their hands for a variety of different reasons. Signers use directed gaze for linguistic functions not found for gaze in spoken languages (e.g., marking role shift or pronominal reference). However, signers are also likely to use gaze as a cue to establish joint visual attention; a function akin to what Gullberg and Holmqvist (1999, 2006) found with co-speech gestures. The fact that both co-speech gesture perceivers and sign perceivers shift fixation from the face toward the hands when their interlocutor is also looking at his or her hands indicates the importance of gaze following and joint attention in linguistic, social interactions.

There was little evidence that the complexity or semantic content of the signed utterance caused a shift in gaze fixation. Although gaze shifts occurred most frequently when the signer was producing a classifier construction compared to a lexical sign or

---

**Table 3** Mean number of gaze shifts toward different types of signs

<table>
<thead>
<tr>
<th>Paint Story</th>
<th></th>
<th></th>
<th>Town Description</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Classifier sign</td>
<td>Fingerspelled word</td>
<td>Lexical sign</td>
<td>Classifier sign</td>
<td>Fingerspelled word</td>
</tr>
<tr>
<td>Native signers</td>
<td>1.71 (1.98)</td>
<td>0.0 (0.0)</td>
<td>0.86 (1.22)</td>
<td>5.0 (2.83)</td>
<td>0.14 (0.38)</td>
</tr>
<tr>
<td>Beginning signers</td>
<td>2.20 (2.25)</td>
<td>0.30 (0.68)</td>
<td>1.20 (1.23)</td>
<td>10.2 (6.39)</td>
<td>1.20 (1.14)</td>
</tr>
<tr>
<td>Total</td>
<td>2.0 (2.09)</td>
<td>0.18 (0.53)</td>
<td>1.06 (1.20)</td>
<td>8.06 (5.74)</td>
<td>0.76 (1.03)</td>
</tr>
</tbody>
</table>

*Note. Standard deviations are in parentheses.*
a fingerspelled word (see Table 3), the shift in gaze appears to be due to the fact that the signer was looking at his or her hands, rather than to the complexity of the classifier construction or the semantic information expressed by the construction. A sign narrator may gaze toward a classifier construction for a number of reasons, for example, for emphasis, to draw the addressee’s attention to a spatial location, or perhaps to linguistically mark object or locative information. Previously, we showed that while producing agreeing or spatial verbs, ASL signers use their eye gaze to mark locations in space associated with grammatical arguments (e.g., objects and locatives; Thompson, Emmorey & Kluender, 2006), and we are currently investigating how such gaze might be used during the production of classifier constructions. For the sign perceiver, however, there were no indications that the complexity of the lexical form or classifier construction caused a shift in gaze away from the face. Overall, gaze fixation patterns during sign language comprehension do not appear to be especially sensitive to variations in linguistic complexity and are unlikely to provide a particularly sensitive measure of processing load.

Funding

National Science Foundation, Linguistics Program (BCS 0517994 to K.E.).

Note

1. Time constraints prevented two deaf participants from taking the ASL narrative comprehension test.

References


Received June 20, 2008; revisions received August 29, 2008; accepted September 3, 2008.