Title
Deaf signers allocate gaze based on type and familiarity of signed input

Journal
Proceedings of the Annual Meeting of the Cognitive Science Society, 46(0)

Authors
Gappmayr, Paris
Lieberman, Amy

Publication Date
2024

Peer reviewed

Deaf signers allocate gaze based on type and familiarity of signed input

Paris Gappmayr (parisgap@bu.edu)
Sargent College of Health & Rehabilitation Sciences, Boston University,
635 Commonwealth Avenue, Boston, MA 02215 USA

Amy M. Lieberman (alieber@bu.edu)
Wheelock College of Education & Human Development, Boston University,
2 Silber Way, Boston, MA 02215 USA

Abstract

In sign languages, gazing towards one's interlocutor is necessary to perceive the language visually. Proficient signers have been found to look at their interlocutor's face, rather than hands, while communicating in ASL. We investigated signers' looks to the face vs. hands while perceiving ASL signs, fingerspelled words, pseudo signs, and fingerspelled pseudowords. Participants' gaze was monitored as they viewed a picture followed by a short, isolated video clip of the corresponding sign or fingerspelled word. We found that participants tended to look at the face more than the hands when perceiving signs vs. pseudosigns, and when perceiving signs vs. fingerspelled words. Age of acquisition did not significantly impact gaze patterns. Results suggest that sign perceivers actively adjust their allocation of gaze based on the perceptual demands of the input.

Keywords: eyetracking, visual attention, sign language perception, deafness, fingerspelling

Background

Sign languages are perceived through the visual modality, and sign comprehension requires active visual attention. In order to perceive signed input, individuals must direct gaze towards their interlocutor. If the sign perceiver is not visually attending to the signer, they will likely miss parts of the signed message, and communication may falter.

Typically, interlocutors will direct gaze to the signer's face in order to perceive the signed input. However, in certain conditions, accurate sign perception may require more purposeful focus on the hands to perceive more fine-grained details of the signed input. This requirement might lead to dynamic shifts in the focus of gaze over the course of a conversational turn. The ways in which signers decide (consciously or unconsciously) to shift gaze to specific focus areas, and the types of signed input that lead to such shifts, have not been well-studied and are not fully understood. The current study aims to address this gap by systematically varying the type and familiarity of input provided to signers while monitoring their gaze during sign perception.

Types of signed input

Lexical Signs

Lexical signs are signs that have conventional form. Signs are described as composed of at least five phonological parameters: Handshape, location, movement, palm orientation, and non-manual markers (Klima & Bellugi, 1979). A lexical sign contains a unique combination of each of these parameters and manipulating one can change the meaning of the sign. Lexical signs are produced using one or two hands, and may be produced anywhere from the neutral space in front of the body to contact on the chest, face, or hands.

Fingerspelling

Fingerspelling in ASL is typically produced on one hand using a unique handshape to represent each of the 26 letters of the English alphabet. Fingerspelling comprises up to one third of spontaneous, interactive signed conversation (Keane & Brentari, 2016), and is used to convey proper nouns, English words which have no direct ASL translation, loan words, and for emphasis. Fingerspelled words tend to be produced at shoulder height in front of the body in the neutral signed space (Keane & Brentari, 2016), which contrasts with the location of lexical signs. Research on fingerspelling perception is scant. Leannah, Wills, & Quandt (2022) studied fingerspelling perception using dynamic Point Light Displays and found that signer proficiency predicted fingerspelling comprehension, and that real words were perceived more accurately than pseudo names.

Schotter, Johnson & Lieberman (2022) presented deaf signers with real and pseudo signs and fingerspelled words at both near and far eccentricities relative to the signers' face. Participants were more accurate in perceiving the signed handshape and fingerspelled letters in the periphery when those stimuli were embedded in real signs and words vs. pseudo signs and pseudo words. This demonstrates that there are higher demands on perceptual abilities when information is presented in the periphery and when it includes unfamiliar stimuli.

Gaze pattern research

Signers are known to look at the face during signed interactions. Emmorey et al. (2009) used head-mounted eye trackers to compare gaze patterns between deaf and hearing ASL signers watching a signed narrative. They found that, while both groups tended to look towards the face, deaf signers tended to look at the eyes, and hearing signers tended to look at the mouth. Mastrantuono et al. (2017) confirmed

In L. K. Samuelson, S. L. Frank, M. Toneva, A. Mackey, & E. Hazeltine (Eds.), *Proceedings of the 46th Annual Conference of the Cognitive Science Society.*

that hearing signers viewing a video of Spanish sign language (LSE) looked towards the face, but found that deaf participants tended to look at the mouth; they hypothesized that, since the majority of their participants had cochlear implants, they were accustomed to looking at the mouth in their daily life for lipreading purposes. They also suggest that viewing modality (video vs. live presentation) impacts looking patterns, resulting in signers looking at the eyes during live, in-person interactions, but at the mouth during video recordings (Mastrantuono et al., 2017). This difference based on viewing context suggests that looking to the eyes is done for social connection, rather than purely linguistic reasons.

Another way that gaze might be socially motivated during signed interactions is by the sign perceiver following the gaze of the sign producer. Emmorey, Thompson, & Colvin (2009) hypothesize that the sign producer looking at their own hands, which is a mechanism sometimes employed by signers in certain contexts, may be a reason why sign perceivers direct gaze to the signer's hands when producing classifier constructions (morphological systems that use handshapes to– often iconically– depict representations of event and object shapes, relative locations, and movements (Sandler & Lillo-Martin, 2006)).

Focusing on the face allows signers to capture the most information when perceiving signed input, with detailed facial expressions in foveal view and larger handshapes perceived in the peripheral vision. However, there is a tradeoff: if sign perceivers focus on the face, they may miss fine-grained details that occur in the handshapes of specific classes of signs, such as fingerspelling. Notably, deaf individuals who use sign language have enhanced peripheral vision relative to hearing individuals (Codina et al., 2017), and can more easily capture movement (Dye et al., 2009; Bosworth & Dobkins, 2002) and handshapes (Emmorey, Bosworth et al., 2009) in the periphery. This may allow for strategic gazing towards the face to capture facial expressions conveyed there, while also allowing for information captured from the hands in the periphery.

Our goal in the current study is to explore the conditions that might lead signers to shift gaze away from the natural focal point of the face and towards the hands when perceiving signed input. Specifically, we manipulated stimuli along two dimensions, comparing lexical signs to fingerspelling, and real ASL signs– which we expected to be familiar to the participants– with phonologically possible but unattested pseudosigns, which were novel to the participants. In addition to sign type and familiarity, we asked whether the age of acquisition of ASL impacts gaze patterns.

We predicted that participants may fixate more on the hands (and hence, less on the face) when perceiving fingerspelling due to its positioning at the outer edge of foveal vision and more complex/rapidly changing handshapes. We also predicted that signers would be more likely to fixate on the hands when perceiving novel stimuli due to having a lack of context and background knowledge for the incoming input. We predicted that there might be an interaction effect, wherein the effect of real vs. novel stimuli would be greater for fingerspelled words than for signs, due to the interaction of a lack of context and increased phonological complexity. Finally, we hypothesized that signers who acquired ASL later in life would fixate more on the hands, particularly when perceiving novel fingerspelled words.

Methods

We used a two-by-two design to compare the effects of sign type (sign vs. fingerspelling) and familiarity (real/familiar vs. novel/pseudosign). Participants completed a picture-sign labeling paradigm, so that context for the familiarity of the upcoming sign could be provided by the preceding picture.

Stimuli Design & Creation

Pictures that served as prompts were colored animation-style clip-art images of either a familiar, recognizable object or an unfamiliar (fake) object, taken from the Novel Object and Unusual Name (NOUN) Database (Horst & Hout, 2016). Stimuli videos consisted of a signer producing either the sign or fingerspelled word that matched the prompt picture. For familiar trials, this was the corresponding canonical ASL sign or fingerspelled word, which were all concrete nouns. For unfamiliar trials, this was either an invented but phonologically possible pseudosign (adapted from Caselli et al., 2021), or an invented but phonologically possible fingerspelled word, taken from the NOUN Database (Horst & Hout, 2016).

Videos were produced by a deaf fluent ASL signer and filmed against a neutral background. Signs (both familiar and novel) were produced below the signer's shoulder to allow the hands and face to be in maximally distinct locations (i.e., none of the selected signs involved locations on the face or above the shoulders). All fingerspelled words, which were all between three to six letters long, were produced in neutral space in front of the signer between their shoulders and waist. To ensure that mouthing did not serve as a confounding variable for gazes to the face, the signer used a neutral facial expression without any mouth movements. Each video began with the signer standing in a neutral position, then raising his hands into frame to produce the token, and lowering his hands again. The signer maintained eye contact with the camera throughout the video clip.

Experiment Design

Prompt pictures and stimulus videos were paired such that each of the sixteen prompt pictures appeared twice: once as a prompt for a sign (familiar or novel), and a corresponding fingerspelled word (also familiar or novel). There were eight trials for each of the four conditions (sign vs. fingerspelling, familiar vs. novel), resulting in thirty-two trials that were presented in one of two randomized possible orders to participants. An example of the trial sequence and conditions is presented in Figure 1.

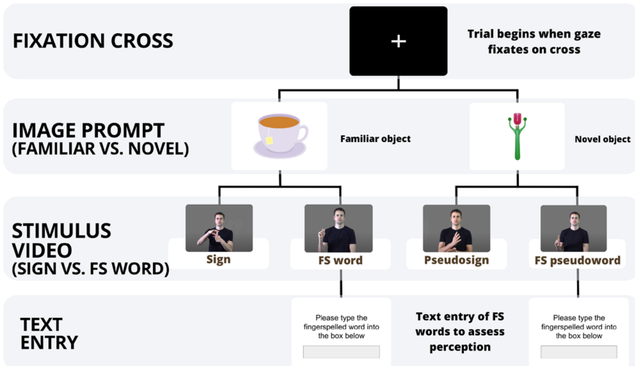

Figure 1: Experiment trial sequence by condition

This figure shows the sequence of events contained within a single trial, arranged by each of the four conditions denoted by the stimulus video. Each trial began with a fixation cross followed by the image prompt and target stimulus video. The task was simply to view the signs, though for fingerspelled trials, participants also entered the text of what they saw.

Participants

Participants were 32 signing deaf & hard-of-hearing (HH) adults between the ages of 18-54 years (mean = 29.6, SD = 11). Participants self-reported their auditory status and the age at which they became deaf/HH: Twenty were identified as deaf and two as HH at birth; seven became deaf between the ages of 1-4, and three became deaf between the ages of 9-12.

All identified ASL as their current primary language. Participants provided a self-rated fluency score on a seven-point scale where one was “I am fluent in ASL” and seven was “I do not use ASL”. Fluency ranged from 1 to 4 (mean = 1.4, SD = 0.83). Twenty-five participants were exposed to ASL from birth to age 5 (henceforth referred to as “early signers”), and seven were exposed between the ages of 5 and 20 (referred to as “late signers”). Eighteen were female, fourteen were male, and four were non-binary. Twenty-five identified their race as White, six as African-American/Black, one as Asian, two as more than one race, and three did not disclose their race. Participants were recruited and tested in the Northeast & Mid-Atlantic US, and all had completed at least some college.

Procedure

Participants were seated in front of a 24-inch LCD display monitor. Their gaze was recorded using an Eyelink 1000 Plus eyetracking camera in remote desktop mode, capturing eye movements at 500 Hz. Following a standard five-point calibration procedure, participants viewed an instruction video and had two practice rounds before beginning the experiment. Each trial began with a gaze-contingent fixation cross. Once the participant directed gaze to the cross, the trial was initiated, wherein a picture was displayed for 500 ms, followed by the stimulus video. In the sign conditions, the trial ended following the presentation of the video. In the fingerspelling conditions, the participant was prompted to type in the word or pseudoword that was displayed in the video to conclude the trial. All stimuli were centered on the screen. Pictures were presented on a white background square measuring 900 x 900 pixels; videos were presented at 1080 x 920 pixels.

Analysis

Eye tracking data during the sign video were analyzed to determine the proportion of time signers spent looking at the hands vs. face. Two interest areas were established, one capturing the head/face, and one on the torso, identifying the region from the shoulders to the waist of the signer to capture the space where the hands signed. Because the signed stimuli did not involve locations on the face, the signer’s hands never moved above the shoulders, allowing for clear separation between the face and the hands.

Data was processed and exported using DataViewer software to obtain information about fixations to the interest areas across each individual trial. Since there were only two interest areas, the data were filtered so that only the trials where more than 70% of the fixations fell within either interest area were analyzed, excluding looks that were outside those regions (28 trials removed). This left a dataset of 995 individual trials. Fixations were then calculated as proportions of looks to the face vs. looks to the hands.

Results

We analyzed fixations using a linear mixed-effects model using the R package ‘lmer’ (from the ‘lme4’ library; Bates et al., 2015), with log-transformed proportion of fixations to the face as the outcome measure, and sign familiarity and sign type as predictors, and random effects for participant (Table 1). We hypothesized that there would be a higher proportion of fixations to the face for familiar, signed lexical stimuli, as opposed to novel and fingerspelled. Our analysis confirmed these predictions: sign type and familiarity independently predicted fixations to the face (Table 1). The proportion of fixations spent looking at the face for a familiar ASL sign perceived by an early signer was 82%. Perceiving a fingerspelled word was associated with a 37% decrease in time spent looking at the face, and perceiving novel stimuli was associated with a 12% decrease. Age of sign acquisition did not significantly predict gaze patterns.

To visualize looking across conditions, we plotted looks in each condition as a difference score, where difference = [proportion of looks to the face] - [proportion of looks to the hands] (Figure 2). The plot reveals an additive effect in which both novel stimuli and fingerspelling lead to increased looks to the hands, such that real lexical signs have the highest proportion of fixations to the face, whereas fingerspelled pseudowords have the lowest proportion of fixations to the face.

Table 1: Predicted gaze to face by condition

	Total Dwell Time Face		
Predictors	*Est.*	*CI*	*p*
(Intercept)	0.82	0.77 – 0.87	**<0.001**
Sign Type [FS]	-0.38	-0.40 – -0.35	**<0.001**
Familiarity [FAKE]	-0.12	-0.15 – -0.10	**<0.001**
AoA [>5]	0.02	-0.09 – 0.13	0.740
Sign Type [FS] × Familiarity [FAKE]	0.02	-0.02 – 0.05	0.344
Random Effects			
σ^2		0.02	
τ_{00} Participant.ID		0.02	
ICC		0.44	
N Participant.ID		32	
Observations		995	
Marginal R^2 / Conditional R^2		0.490 / 0.714	

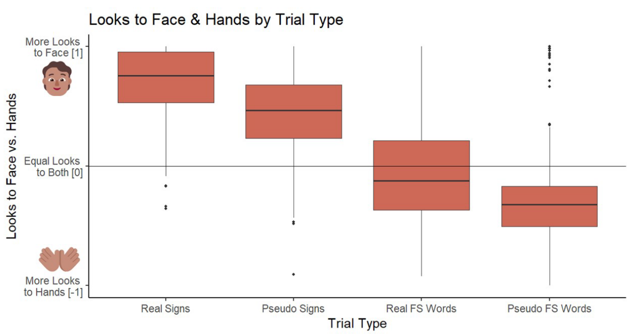

Figure 2: Looks to face and hands by trial type.

This figure shows the proportion of looks to the face vs. hands for each of the four conditions represented on the x axis. On the y axis, a 0 indicates equal looks between the face and hands, while positive values indicate greater proportion of fixations to the face, and negative values indicate greater proportion of fixations to the hands.

Contrary to our expectations, the interaction between sign familiarity and sign type did not have a significant impact on time spent looking to the hands: signers shifted their gaze to the hands in similar proportions in both the signed and fingerspelled conditions regardless of the familiarity of the stimuli.

We speculated that one reason we didn't see the interaction effect we expected is because signers acclimated to the various conditions over the course of the experiment. The signer produced signs with no marked facial expression and no mouth movements. If participants realized in the initial trials that the signer's facial expression did not contain any linguistic information and was not a reliable cue to sign identity, they may have then directed fewer looks to the face over the course of the experiment. To examine this possibility, we re-ran the model and added trial index (each individual trial's order in the sequence of the experiment) as a predictor. Trial index was a significant predictor: each increase in trial index was associated with a slight decrease in looks to the face, so that by the final trial, looks to the face decreased by 10% compared to the first trial (Est: -0.003, Std. Err: 0.0007, df: 958, t: -4.873, $p < 0.001$). The effect of trial index was independent of sign condition- neither interaction between trial index and sign type or familiarity were significant (ns, $p > 0.1$).

Discussion

This study investigates the relationship between looks to the face vs. hands while perceiving lexical signs or fingerspelled words. Our research question is to determine the proportion of time signers spend fixating on the face vs. hands while perceiving lexicalized signs vs. fingerspelled words that are either novel or familiar, and if this varies as a function of the age of ASL acquisition.

Effect of Sign Type

We found an effect of sign type- participants looked to the hands more during fingerspelling, regardless of familiarity. This was expected, as fingerspelling involves more detailed fine motor changes to handshape that could make perception in peripheral vision more difficult. The experiment was designed to increase task demand by requesting participants type back the fingerspelled word they saw. Thus, it was advantageous for them to look at the hands to ensure total accuracy. Existing research shows that signers perceiving fingerspelled words via point light displays were less accurate perceiving novel names than familiar names (Leannah et al., 2022). In the case of our study's pseudoword condition, upcoming letters couldn't be predicted from context, so there were additional benefits to focusing gaze on the hands to correctly perceive the pseudoword.

Effect of Familiarity

In addition to sign type, we also manipulated the familiarity of the stimuli, revealing a significant effect of familiarity. Our results suggest that in situations where sign perceivers have context and can therefore reasonably predict what the upcoming sign or fingerspelled word will be, they might fixate continuously on the face and use their peripheral vision to confirm their prediction about the sign or fingerspelled word. However, when perceiving novel stimuli, there is no model against which they can map the input, and peripheral perception is likely not enough to capture the fine-grained details of the phonological features (Schotter et al., 2020). Therefore, sign perceivers direct gaze away from the face to the hands, possibly in order to more accurately perceive the stimuli.

The Role of Age of Acquisition

Previous research has suggested that gaze patterns might vary during sign perception depending on the signer's hearing status and fluency (Emmorey et al., 2009; Mastrantuano et al., 2017). Following this line of thought, we hypothesized that signers with later access to ASL might be more likely to shift gaze to the hands, possibly due to more experience allocating gaze to best capture the visual input from an earlier age. However, we did not find a significant difference in the proportion of fixations to the face vs. hands signers exposed to ASL before vs. after early childhood. We consider several possible interpretations. First, the groups were not evenly split; there were far more "early signers" than "late signers". The uneven distribution between these groups may have masked any possible effect of AoA. Alternatively, given that hearing signers allocate gaze towards the face (Emmorey et al., 2009), AoA may not be a reliable predictor of gaze patterns during sign perception. Rather than having some sort of critical acquisition period, gaze may be more malleable and experience-based. Given that all participants in this sample were skilled signers who used ASL frequently, they may have had adequate experience to develop robust gaze patterns.

Limitations & Future Research

Importantly, the current sample includes all deaf/HH signers. Deaf individuals are known to have advanced peripheral perception for certain types of stimuli (Chen et al., 2006; Nava et al., 2008; Dye et al., 2009), which could impact their gaze patterns during sign perception. To tease apart the effect of hearing status and sign proficiency on gaze patterns, in ongoing data collection we are including hearing signers as an additional participant group. By looking at a wider range of signing abilities, we will be able to tease apart the effects of proficiency and experience from age of acquisition on gaze during sign perception.

Additionally, we plan to run another version of this experiment with stimuli that include mouthing and non-manual markers, to compare gaze patterns when there is no linguistic information on the face to more naturalistic conditions. This will allow us to determine how much of gaze behavior is adaptive to match the input and situation, while also providing conditions that better match those of real-world signing interactions.

Finally, given that the current study involved perception of signs on a monitor, it is unclear if these results would generalize to live, in-person interactions. The screen-based viewing context of the experiment meant there was no social expectation to maintain eye contact, further allowing for gaze to be shifted towards the hands. Further investigation in a live interaction setting will allow us to explore how social interactions impact gaze patterns during signed conversation.

Conclusion

Signers – whether children or adults – need to learn to allocate their gaze as an important step in the acquisition of language and social cognition. In order to understand the trajectory of the development of gaze control, we first need to know how skilled signers control their gaze. The current study highlights the dynamic nature of gaze allocation during sign perception. Although ordinarily sign perceivers receive lots of linguistic and social information from the face, in situations where they don't have context for the topic or lexical item, or when perceptual demands are high, signers appear to shift gaze to the hands. When there is both a lack of context and high perceptual demand (i.e. perceiving fingerspelled pseudowords), the effect is additive. Signers efficiently allocate attention dynamically in response to the perceptual demands of the input.

Understanding the gaze allocation of skilled signers can serve as a goalpost for skills that children and new signers ought to acquire and can have important implications for education of both deaf/HH children and hearing adults who are acquiring ASL as an L2 later in life.

References

Agrafiotis, D., Canagarajah, N., Bull, D. R., & Dye, M. (2003). Perceptually optimised sign language video coding based on eye tracking analysis. Electron. Lett. 39, 1703–1705. doi: 10.1049/el:20031140

Argyle, M., & Cook, M. (1976). Gaze and Mutual Gaze. Oxford: Cambridge University Press.

Bates D, Mächler M, Bolker B, & Walker S (2015). "Fitting Linear Mixed-Effects Models Using lme4." Journal of Statistical Software, 67(1), 1–48. doi:10.18637/jss.v067.i01.

Bosworth R. G., & Dobkins K. R. (2002). The effects of spatial attention on motion processing in deaf signers, hearing signers, and hearing nonsigners. Brain Cogn. 49, 152–169. 10.1006/brcg.2001.1497

Caselli, N. K., Emmorey, K., & Cohen-Goldberg, A. M. (2021). The signed mental lexicon: Effects of phonological neighborhood density, iconicity, and childhood language experience. Retrieved from osf.io/8a7tb

Codina, C. J., Pascalis, O., Baseler, H. A., Levine, A. T., & Buckley, D. (2017). Peripheral Visual Reaction Time Is Faster in Deaf Adults and British Sign Language Interpreters than in Hearing Adults. Frontiers in psychology, 8, 50. https://doi.org/10.3389/fpsyg.2017.00050

Emmorey, K., Thompson, R., & Colvin, R. (2009). Eye Gaze During Comprehension of American Sign Language by Native and Beginning Signers. The Journal of Deaf Studies and Deaf Education. 14(2). Pages 237–243, https://doi.org/10.1093/deafed/enn037

Dye, M. W. G., Hauser, P. C., & Bavelier, D. (2009). Is visual selective attention in deaf individuals enhanced or deficient? The case of the useful field of view. PloS One, 4(5), e5640–e5640. https://doi.org/10.1371/journal.pone.0005640

Horst, J. S., & Hout, M. C. (2016). The Novel Object and Unusual Name (NOUN) Database: A collection of novel images for use in experimental research. Behavior research methods, 48(4), 1393-1409.

Keane J, & Brentari D. (2016). Fingerspelling: Beyond handshape sequences. The Oxford handbook of deaf

studies in language.;146–60. https://doi.org/10.1093/oxfordhb/9780190241414.013.10
Leannah, C., Willis, A. S., & Quandt, L. C. (2022). Perceiving fingerspelling via point-light displays: The stimulus and the perceiver both matter. PloS one, 17(8), e0272838. https://doi.org/10.1371/journal.pone.0272838
Mastrantuono E, Saldaña D & Rodríguez-Ortiz IR (2017) An Eye Tracking Study on the Perception and Comprehension of Unimodal and Bimodal Linguistic Inputs by Deaf Adolescents. Front. Psychol. 8:1044. doi: 10.3389/fpsyg.2017.01044
Muir, L. J., & Richardson, I. E. G. (2002). "Video telephony for the deaf: analysis and development of an optimised video compression product," in Proceedings of the Tenth ACM International Conference on Multimedia (New York, NY: ACM), 650–652. doi: 10.1145/641007.641137
Sandler, W. & Lillo-Martin, D. (2006). Sign Language and Linguistic Universals. Cambridge University Press. ISBN 978-0521483957
Schotter, E. R., Johnson, E., & Lieberman, A. M. (2020). The sign superiority effect: Lexical status facilitates peripheral handshape identification for deaf signers. Journal of Experimental Psychology: Human Perception and Performance, 46(11), 1397–1410. https://doi.org/10.1037/xhp0000862