Title
Using transitional information in sign and gesture perception.

Authors
Brozdowski, Chris
Emmorey, Karen

Publication Date
2023-06-01

DOI
10.1016/j.actpsy.2023.103923

HHS Public Access
Author manuscript
Acta Psychol (Amst). Author manuscript; available in PMC 2023 October 14.

Published in final edited form as:
Acta Psychol (Amst). 2023 June ; 236: 103923. doi:10.1016/j.actpsy.2023.103923.

Using transitional information in sign and gesture perception

Chris Brozdowski[*], **Karen Emmorey**[*]
San Diego State University, University of California, San Diego, United States of America

Abstract

For sign languages, transitional movements of the hands are fully visible and may be used to predict upcoming linguistic input. We investigated whether and how deaf signers and hearing nonsigners use transitional information to detect a target item in a string of either pseudosigns or grooming gestures, as well as whether motor imagery ability was related to this skill. Transitional information between items was either intact (Normal videos), digitally altered such that the hands were selectively blurred (Blurred videos), or edited to only show the frame prior to the transition which was frozen for the entire transition period, removing all transitional information (Static videos). For both pseudosigns and gestures, signers and nonsigners had faster target detection times for Blurred than Static videos, indicating similar use of movement transition cues. For linguistic stimuli (pseudosigns), only signers made use of transitional handshape information, as evidenced by faster target detection times for Normal than Blurred videos. This result indicates that signers can use their linguistic knowledge to interpret transitional handshapes to predict the upcoming signal. Signers and nonsigners did not differ in motor imagery abilities, but only non-signers exhibited evidence of using motor imagery as a prediction strategy. Overall, these results suggest that signers use transitional movement and handshape cues to facilitate sign recognition.

Keywords

Transitions; Prediction; American Sign Language; Pseudosign; Gesture

1. Introduction

In spoken language, transitions between segments may be very short as the vocal articulators are small and move rapidly in a restricted space. In addition, the structure of the mouth hides many transitional movements. Evidence of transitional movements is primarily observed through the blending of phonological features across neighboring segments in the form of coarticulation across syllables and words (see Recasens, 2018 for review). The articulators of sign language, by contrast, are consistently present and must move larger distances from the offset of one monosyllabic sign to the onset of the next (signs tend to be monosyllabic; Brentari, 1998). Sign languages do exhibit coarticulations within signs (e.g., Mauk, 2003),

[*]Corresponding authors at: Laboratory for Language and Cognitive Neuroscience, 6495 Alvarado Road, Suite 200, San Diego, CA 92120, United States of America. cbrozdowski@yahoo.com (C. Brozdowski), kemmorey@sdsu.edu (K. Emmorey).

Declaration of competing interest
The authors report no conflict of interest.

but here we focus on the perception of transitional periods between signs. Native signers have consistent experience monitoring highly visible articulators as they transition from one sign to the next. We ask what role different types of information within these transitions may play in sign language comprehension.

As a means of establishing that these periods are distinct from the signs themselves, Jantunen (2013) investigated the physical characteristics of these periods during natural signing. According to this work, the transition period begins when the dominant hand diverts from the lexical movement specified for the sign; the transitional period ends immediately prior to the specified movement for the subsequent sign. Using this demarcation method, Jantunen (2013) examined fluid narrative signing and found differences between signs and transitions in both speed and acceleration profiles. Lexical movements, in contrast to transitional movements, have reduced velocity and increased acceleration. Viewed from the lens of communicative intent, signers may be slowing down to better communicate intentional information and increasing acceleration to more rapidly reach the desired movement target of the sign. This physical contrast between lexical movement and the interim movement between signs may facilitate perceptual parsing of intentional vs. transitional movements.

Klima et al. (1999) provide evidence that sign language experience supports the ability to parse intentional from transitional movements. Researchers filmed a native Chinese speaker drawing pseudocharacters (plausible but nonexistent Chinese characters) in the air, with a light attached to their fingertip in an otherwise dark environment. Signers and nonsigners were asked to write down the pseudocharacters observed in this point-light display. Both Chinese and American signers showed improved ability to dissect the continuous fingertip movement into intentional and transitional movements, such that they included fewer transitional movements in their character drawings and lifted their pen more often than their nonsigner counterparts. Thus, signers both produce (Jantunen, 2013) and perceive (Klima et al., 1999) a contrast between intentional and transitional movements, validating this distinction.

However, Jantunen (2015) has argued against this distinction and proposed that the onset of a sign be defined as the first detectable articulatory feature in the visual signal (e.g., the handshape). Critically, such a feature may appear as the signer's hand moves toward a target location on the body or in signing space. Under this view, there is minimal or no linguistically relevant information within a transition because the movement phase that would be traditionally considered transitional contains the information that defines the beginning of the sign. Further, for signs in context (e.g., in sign language corpora), "there are no real transitions between signs" (Jantunen, 2015; p. 120). Here, we adopt the more traditional definition of sign onset (e.g., when the hand contacts or moves away from a target location on the body) and consider the preceding and following movements as transitional (for details see Crasborn et al., 2015; Johnson & Liddell, 2011; Caselli et al., 2017; Emmorey et al., 2022). This view maintains a clearer distinction between phonologically-specified and transitional movements for signs and between preparatory and stroke (intentional) movements for gestures (Kita et al., 1998).

The utility of transitions for comprehension may differ with respect to the phonological information that is present simultaneously in the transitional period, in particular the different phonological units or parameters of sign: handshape, location, and movement. Across these parameters, movement might drive recognition of the onset of the next sign. Specifically, the direction and velocity of movement would allow a signer to infer the onset location of a subsequent sign. For example, assuming a low resting position in the lap, a hand moving upward at high velocity could probabilistically give information about a sign's onset location at the forehead. A slower velocity might signal an onset location as the chest. By narrowing the possibility space for the subsequent sign location, a signer would be more rapidly able to activate the relevant sign. While there is significant evidence for coarticulation effects for handshape and location in sign production (Cheek, 2001; Mauk, 2003; Ormel et al., 2013, 2017), evidence that comprehenders can make use of coarticulation has focused primarily on location. For example, Grosvald and Corina (2012) demonstrate that, with explicit instruction, both signers and nonsigners are capable of detecting anticipatory location co-articulation. To our knowledge, no study has specifically examined the utility of movement transitions (e.g., exploiting movement velocity), but there is evidence that larger movements provide more information.

Arendsen et al. (2007) examined sign detection times for both signers and nonsigners by presenting participants with sequences of manual "fidgets" that also included one sign in Sign Language of the Netherlands (NGT). The target sign was either preceded by either hands resting on a table or non-communicative fidgeting (e.g., scratching the nose, running hands through hair). Notably, all stimuli included the transitional movement from either rest or from the preceding fidget. Participants were instructed to discriminate between signs and fidgets by pressing a key as soon as they identified the sign. Signers could draw on their knowledge of NGT, while the nonsigners made this judgment purely based on gesture form (with minimal training).

On average, all participants identified the sign 200 ms after sign onset, but signers were 90–150 ms faster than nonsigners at identifying signs, which is indicative of language experience giving rise to richer or more efficient predictive representations of incoming information. Importantly, across all participants, there was a negative correlation between the duration of the transitional movement and response time to detect the sign. In other words, the longer the sign model took to transition, the quicker participants identified a given item as a sign. In the case of the signers, transitional movements likely informed the linguistic process of sign identification whereas for nonsigners the transitional movements likely provided cues to form differences between signs and fidgets. It is not clear, however, to what extent this improved prediction was based solely on longer transitional movements or on extended exposure to handshape coarticulation during the transition.

Electrophysiological research is well suited to look at the time course of processing coarticulation effects for different phonological features during sentence comprehension. In an event-related potential (ERP) study, Hosemann et al. (2013) examined the N400 component, which indexes lexico-semantic processes (see Kutas & Federmeier, 2011, for review) in response to German Sign Language (DGS) sentences that contained either semantically predictable or unpredictable endings. For example, *a rabbit jumps across a* *path* is much more predictable than *a crocodile jumps across some meat*. Per DGS syntax, JUMP is the last sign in each case, and if the preceding nouns are CROCODILE and MEAT, EAT is the expected verb and JUMP is anomalous. The N400 is typically observed 300–500 ms after the onset of a semantically anomalous word. In the case of Hosemann et al. (2013), N400 effects (larger negativities for the anomalous vs. expected sign) were less clear when timelocked to sign onset, but more typical when timelocked to the handshape change (i.e., when the target handshape was formed prior to sign onset). In other words, the transitional handshape, and not sign onset, drove early indicators of comprehension. Signers are therefore using transitional handshapes to support sign recognition (see also Emmorey et al., 2022).

Further evidence of signers attending to intermediate handshapes can be seen in studies of fingerspelling, the rapid sequential presentation of sign handshapes that correspond to letters in an alphabetic orthography. Kinematic studies indicate that fingerspelled words are not comprised of a series of static handshapes, but rather there are dynamic, co-articulatory movements between handshapes that create an overall shape or *envelope* for a fingerspelled word (Keane & Brentari, 2016; Wilcox, 1992). Geer and Keane (2014) argue that fingerspelling comprehension relies more on the overall envelope which includes transitional handshapes, rather than sequential processing of the individual letter handshapes. To assess this hypothesis, Geer and Keane (2017) provided students with either (a) standard classroom exercises to reinforce the fingerspelling alphabet from American Sign Language (ASL), or (b) explicit training in the contrast between transitions and canonical fingerspelling segments, as well phonetic variation (i.e., how a letter may look noncanonical in context). Only the students with training in identifying transitional information showed improvement in fingerspelling comprehension. We suggest that native signers are already taking advantage of these transitional handshape cues for sign comprehension.

The present experiment examines whether or not signers and nonsigners differ in their ability to incorporate transitional handshapes into predictive representations, over and above movement information like velocity and trajectory. We also ask whether or not the ability to attend to transitional handshape is confined to language processing or also occurs for non-linguistic actions, such as fidgets or grooming gestures.

To examine the features that drive predictive representations, we selectively degraded transitional information in different ways, while preserving the timing of presentation, in a response-to-target task. *Normal* videos presented fluent strings of either pseudosigns (possible but non-occurring ASL signs) or grooming gestures (such as rubbing the eyes) with no alternations. *Blur* videos selectively blurred the hands during transitions. Blurring the hand prevents the use of intermediate handshapes in facilitating predictions about the upcoming item, while still preserving broader movement information, such as the trajectory of the hand(s). In *static* videos the frame immediately prior to the transition was frozen for the entire transitional period, completely removing all transitional information. Both deaf signers and hearing nonsigners watched these videos and were asked to monitor for a target item that could occur at any point in the string. Our dependent measure was response time (RT) for target detection, and we expected that a stronger predictive representation would result in faster RTs. The different video conditions allowed us to investigate the use of transitional handshape cues (Blur vs. Normal videos) and the use of transitional movement cues (Blur vs. Static videos).

To examine the domain specificity of potential group effects, videos featured both pseudosigns and grooming gestures. For the signers, pseudosigns had phonological structure (e.g., they conformed to the phonological and phonetic constraints of ASL), whereas for nonsigners these forms were simply meaningless gestures. Grooming gestures were non-linguistic self-adjustment actions that an individual might exhibit in day-to-day life (e.g., scratching one's chin). Grooming gestures served as a baseline to which groups had equal exposure, and pseudosigns probed for the effects of sign language experience and knowledge.

Regarding potential group effects, regular communication involves active predictive representations of interlocutors (Pickering & Garrod, 2013). We hypothesized that sign language experience would bolster all manual predictive representations, and therefore we predicted that signers would have overall faster response times compared to the nonsigner group. As previous evidence points to signers making use of transitional movement (Arendsen et al., 2007) and transitional handshapes (Cormier et al., 2008; Hosemann et al., 2013) during language comprehension, we further predicted that signers would show greater RT penalties due to the loss of transitional information (i.e., slower RTs for Blur compared to Normal videos and slower RTs for Static compared to Blur videos). We hypothesized that sign language experience would enhance the ability to incorporate transitional information into predictions, and that this effect would be most pronounced in linguistic contexts.

Finally, one component of the ability to generate predictions from transitional information likely involves more general motor imagery abilities. Motor imagery refers to the ability to mentally simulate actions. While signers undoubtedly have greater experience perceiving and executing linguistically contrastive handshapes and motor movements, it is not yet known whether this experience reshapes their motor imagery abilities. One possibility is that sign language experience is tied to drawing distinctions between contrastive handshapes: detailed fine-motor representations are unimportant so long as one can faithfully associate the perceptual input with the relevant category. Support for this hypothesis comes from studies showing categorical perception for handshapes (Baker et al., 2005; Emmorey et al., 2003; Palmer et al., 2012). On the other hand, and from the perspective of motor simulation proposals (i.e., Pickering & Garrod, 2013), sign comprehension may involve covert motor imitation of not only target handshapes, but transitional handshapes that do not conform to typical linguistic distinctions (Geer & Keane, 2014; Schwarz, 2000). To help understand the role of motor imagery in manual predictions and transition perception. We used the Test of Ability in Movement Imagery for hands (TAMI-h; Donoff et al., 2018) to investigate whether motor imagery ability is related to the ability to predict an upcoming action (sign or gesture), as well as whether sign language experience might enhance hand motor imagery.

2. Methods

2.1. Participants

Forty-two right-handed participants were recruited from the community in San Diego via fliers, online postings, and word of mouth. Twenty-one deaf signers (13 female, *M* age = 35.5, range: 23.9 to 59.3, *SD* = 9.8), who were either native signers (i.e., born into deaf, signing families; n = 10) or were early signers (exposed to ASL before age six; n = 11), and 21 sign-naïve English speakers (12 Female, *M* age = 29.0, range: 19 to 59, *SD* = 11.2) participated in the study. Nonsigners reported no or minimal knowledge of a sign language (e.g., a few signs or the fingerspelled alphabet).

2.2. Materials

Participants saw twelve videos for each of six conditions in a 2 × 3 design: stimulus Type (grooming gesture, pseudosign) by transition Condition (Normal, Blur, Static transitions). Twelve common grooming gestures and 12 pseudosigns from Brozdowski and Emmorey (2020) were used in this experiment. See Fig. 1 for the full inventory of stimuli. Each item involved the dominant hand touching some part of the body, including the non-dominant hand, the nose or elsewhere on the head or face. The pseudosign stimuli were designed to match grooming gestures in both place of articulation and the number of hands involved (half were one-handed and half were two-handed).

Each stimulus video presented a randomized eight-item subset of the twelve-item inventory. A hearing signer model produced eight items as naturally as possible in a fluent string, including fluid transitions from the offset of one item to the onset of the next. Twelve such strings were generated while avoiding the repetition of any individual item within a subset and including each of the twelve items exactly eight times across all strings. Item order and targets were matched across stimulus types (e. g., for grooming gesture string 1, gesture 7 was the target in third position, and, for pseudosign string 1, pseudosign 7 was the target in third position).

2.2.1. Condition—Each of the twelve videos was edited twice to manipulate the availability of transitional information. The Normal videos were presented as filmed. The Blur videos added blur to the hands during transitional periods for the same videos using an ellipse mask layer and a gaussian blur key-framed to fit precisely over the hand in Adobe® Premiere Pro CC (see Fig. 2). Transitional periods were identified using the definition present in Jantunen (2013). The Static condition preserved the final frame of a pseudosign or gesture for the entire transition time, also using Adobe® Premiere Pro CC, which removed all transitional information. To the observer, this manipulation appears to be rapid sequential presentation of stimulus items with a still frame in between. All stimuli were shown at 1920-by-1080 pixel resolution, at 60 frames per second. Importantly, the timing of videos was never manipulated; the delay between the first and second item in any given string was preserved across conditions.

2.2.2. Test of Ability in Motor Imagery for Hands (TAMI-h)—The TAMI-h was designed by Donoff et al. (2018) and includes four distinct versions: two to assess left hand imagery and two to assess right hand imagery. For the present study we only presented the right-hand versions, counterbalanced across participants. Participants sat at a table and were given a packet, pen, answer sheet, and tennis ball. They were instructed to (a) read the imagery instructions and each of the questions carefully, (b) imagine the handshape described by the instructions for each trial, (c) turn the page and do not look back to the questions, and (d) hold the tennis ball in the right hand firmly while both reading and selecting a response. The tennis ball suppressed overt muscle activity. Below is an example of one of the imagery instructions.

1. Lay your hand open, palm up, with your fingers together.
2. Spread your fingers apart.
3. Cross your pinky finger in front of your ring finger.
4. Point your middle finger perpendicular to the palm.
5. Touch the tip of your thumb midway up your middle finger.

(Donoff et al., 2018, p. 4)

Approximately half of the trials were accompanied by a set of pictures of hands, targeting isolated hand imagery; participants match the resulting imagined handshape to pictures of hands. The other half of trials targeted functional hand imagery; participants match the resulting handshape to the appropriate commonplace object. Motor imagery scores were calculated as a percent correct.

2.3. Procedure

Participants first completed the response-to-target task and then took the TAMI-h. For the response-to-target task, each participant sat in front of an iMac running OSX 10.8 and Psyscope X B77 (Cohen, 1993). Instructions and practice items demonstrated the structure of the task. The English instructions are presented below, and these instructions were presented in ASL for the deaf signers.

> The goal today will be to see how quickly you can identify an invented sign or gesture in context. We will look at this by first showing you a picture of a 'target' item. Press the space key to continue to the following video. This video will show several invented signs or gestures in a sequence. Please press the B key as soon as you see the target item. If you forget the picture that was shown, please wait until video ends. If you press the B key and nothing happens, please also wait until the video ends. It's important that you focus and press B as quickly as you can when you see the target item. Sometimes, you will see blurring between invented signs or gestures, or the video will freeze before moving on to the next items. Please do your best to ignore blurring and freezing and only try to predict the target sign or gesture. Let's start with the (pseudosigns/gestures). You will now see a video of each of the possible target items.

The participant then saw an isolated video of each item of the 12 items in the relevant stimulus type's inventory. In each demonstration video, the model moved from rest position (arms relaxed, hands by her sides) to produce the stimulus item, and then moved her hand(s) back to the rest position. The participant then saw eight practice trials. For each trial, the participant saw a still frame from one of the possible target items for that stimulus type. The participant was asked to press a button as soon as they recognized the target within in the video string. Reaction times (RTs) were collected for button-press relative to the onset of the target item.

Each string of eight items was shown four times with different targets, resulting in 48 RTs per condition and 288 data points per participant. Each of the 24 possible items (see Fig. 1) was shown four times for each of three video conditions. The experiment took 30–45 min to complete. Trials were balanced such that the target was equally likely to occur in any of the eight possible positions within a string (first, second, third, etc.); six items were shown for each position within each condition. On average, the target appeared 5.2 s into the video (range: 0.6–11.1 s, $SD = 3.1$ s). Stimulus type (pseudosigns, grooming gestures) was blocked and counterbalanced across participants. At the halfway point within each block, as well as between blocks, participants were prompted to take a break.

Keypresses were recorded using Psyscope (Cohen, 1993). While participants were permitted to respond either before or after the onset of the target item, and RTs were measured relative to this target onset, efforts were taken to remove any false positives or errant key presses from the data. Specifically, responses were labeled too early if they occurred prior to the transition immediately preceding the target item, because a participant cannot identify a target prior to any phonological features being present. Responses were labeled too late if the response time was greater than two standard deviations above the mean for that specific combination of target, video and participant group.

3. Results

Of approximately 11,800 data points, 1032 or 8.7 % were rejected if they did not include a response or if the response was labeled too early or too late. While more data points were rejected for nonsigners than signers (10.4 % vs. 7.2 %, respectively), this difference was not significant, $t(20) = 1.34$, $p = 0.18$. Further analyses of this dataset used linear mixed effects models (LMEs), with post-estimation contrasts of marginal linear predictions. LMEs were conducted with Stata 14.2 (StataCorp, 2015).

To examine the role of video condition in predicting response time, group, video condition, stimulus type, and all interactions were included as fixed factors. Subject, item identity, and target placement (an ordinal variable ranging from one to eight) were included as random intercepts. In general, reaction times improved for later placements (means by placement half: $M_{position\ 1–4} = 572$ ms, $M_{position\ 5–8} = 440$ ms). Finally, both video condition and stimulus type were included as random slopes relative to the subject intercept. In essence, this model focused on a full factorial design of condition (Normal vs. Blur vs. Static), stimulus type (grooming gesture vs. pseudosign), and group (signer vs. nonsigner), while parsing out variation due to early vs. late targets and variation in difficulty across items. Random slopes of conditions relative to subject allow for between-subject variation in conditions of interest (video condition and stimulus type). One participant, for example, showing a huge contrast between grooming gestures and pseudosigns, will not sway the overall analysis. The omnibus test of this model is shown in Table 1 and parameter estimates for each condition are shown in Fig. 3.

Counter to expectations, there was no main effect of group. There was, however, a main effect of condition, with faster RTs for Normal compared to Static conditions, $p <$ 0.001. This result indicates that participants were using transitional information to generate predictions. Next, the two-way interaction between Group and Type, $p = 0.018$, indicates that signers and nonsigners were differentially affected by the type of stimulus. A contrast of marginal linear predictions highlights that, while signers were equally fast to detect grooming gesture and pseudosign targets, $\chi^2(1) = 0.25$, $p = 0.62$ ($M = 497$ ms, $SE = 26.4$ vs. $M = 504$ ms, $SE = 26.4$), nonsigners were significantly faster to detect grooming gesture than pseudosign targets, $\chi^2(1) = 9.44$, $p < 0.01$ ($M = 511$, $SE = 26.4$ vs. $M = 555$, $SE =$ 26.4).

Planned contrasts for this experiment included an examination of whether each group attended to movement (Blur vs. Static) or handshape transition (Normal vs. Blur) information in predicting target signs. Contrasts indicated that both signers, $\chi^2(1) = 120.71$, $p < 0.001$, and nonsigners, χ^2 (1) = 179.2, $p < 0.001$, attended to movement information. This pattern is clearly visible in the steep slopes between Blur and Static conditions in Fig. 3. Both groups show slower RTs for Static compared to Blurred videos in both grooming gesture and pseudosign conditions.

An additional planned contrast, looking at the use of handshape information within a transition revealed that signers showed significantly faster RTs for Normal compared to Blurred videos, $\chi^2(1) = 37.68$, $p < 0.01$, while nonsigners did not show a difference between these conditions, $\chi^2(1) = 2.14$, $p = 0.14$. This effect was driven by the pseudosign stimuli, $\chi^2(1) = 55.50$, $p < 0.01$ ($M = 397$ ms, $SE = 27.5$ vs. $M = 510$ ms, $SE = 27.8$), as signers exhibited no significant difference between the Normal and Blurred videos for grooming gestures, $\chi^2(1) = 2.04$, $p = 0.15$ ($M = 435$ ms, $SE = 27.5$ vs. $M = 456$ ms, $SE = 27.7$). While both groups made use of transitional motion information for both stimulus types, only the signers made use of handshape transitions, and did so only for pseudosigns. Nonsigners did, however, trend in the same direction for pseudosigns (Normal: $M = 484$ ms, $SE = 27.7$, Blur: $M = 524$ ms, $SE = 27.9$).

A post-hoc test revealed that signers were faster at responding to Normal pseudosign stimuli than nonsigners, $\chi^2(1) = 5.04$, $p = 0.02$ ($M = 397$ ms, $SE = 27.5$ vs. $M = 484$ ms, $SE =$ 27.7). Under ordinary video conditions, signers showed faster response-to-target times for pseudosign videos than nonsigners.

This experiment was relatively long, and participants may have had an opportunity to improve their performance and sensitivity to transitional information over the course of the experiment. A few participants anecdotally reported that the task became easier over time. To explore this possibility, we conducted a post-hoc follow-up model that also included a fixed factor for split-half stimulus type condition progress, examining possible habituation effects, or reduction in video condition effects over the course of each stimulus type block. As the previous analysis showed no between-group differences in the use of movement information (Blur vs. Static), the following analysis focused only on handshape transition habituation (Normal vs. Blur). The omnibus test of this model is shown in Table 2 and parameter estimate differences (Blur – Norm) are shown in Fig. 4, where positive values indicate improved response times to videos with handshape information present. Through this model, we can examine how attention to transitional handshape information may have changed over the course of either the pseudosign or grooming gesture conditions. Due to the post-hoc nature of these eight tests, contrasts of marginal linear predictions will be reported relative to a Bonferonni corrected alpha level of $p = 0.006$.

Signers showed significantly faster RTs for the Normal compared to Blurred videos for the pseudosign condition for both the first half, χ^2 (1) = 35.59, $p < 0.0001$, and second half of trials, χ^2 (1) = 20.07, $p < 0.0001$, indicating consistent use of pseudosign handshape transitions throughout the experiment. Signers also showed significantly faster RTs for the Normal compared to Blurred videos in the grooming gesture condition, but only in the first half of trials, χ^2 (1) = 9.44, $p = 0.0021$. Mean differences indicate that, over the course of the experiment, signers improved their target detection for grooming gestures in the Blurred video condition (M_{First} = 474 ms vs. M_{Second} = 440 ms), but their performance decreased for grooming gestures in the Normal video condition (M_{First} = 408 ms vs. M_{Second} = 460 ms). It may be the case that signers developed a strategy that deemphasized or ignored handshape in order to try to perform well in both the Blur and the Normal conditions, and this strategy could have resulted in poorer performance over time in the Normal condition. Additionally, nonsigners showed evidence of attending to pseudosign handshape transitions only during the second half of trials, χ^2 (1) = 13.78, $p = 0.0002$. Means indicate that this effect is driven by improved response times to pseudosigns in the Normal videos (M_{First} = 542 ms vs. M_{Second} = 433 ms), as pseudosign RTs for the Blurred videos were relatively stable over time (M_{First} = 540 ms vs. M_{Second} = 512 ms). This pattern of results indicates that signers start off using handshape transition information for grooming gestures, but this effect disappears in the second half of the experiment. In contrast, nonsigners showed evidence of increased an ability to use handshape transition information for pseudosigns in the second half of the experiment.

3.1. Hand motor imagery results

Analyses of data from the TAMI-h were conducted in SPSS 25 (IBM Corp, 2017). Mean scores for each group for the two subtests (isolated hand imagery and functional imagery) are shown in Fig. 5. First, we examined group by subtest effects in an ANOVA on overall accuracy. There was a marginal effect of subtest, $F(1,51) = 3.251$, $p = 0.077$. Participants had marginally higher scores for isolated hand imagery trials (M = 63 %, SD = 27 %) compared to functional imagery trials (M = 57 %, SD = 16 %). This trend aligns with the significant difference found in Donoff et al. (2018). There was neither a main effect of group, $F(1,51) = 0.151$, $p = 0.70$, nor an interaction between group and subtest, $F(1,51) = 0.281$, $p = 0.60$.

Next, TAMI-h scores and response-to-target RTs were compared in bivariate correlations for each group. While signers showed no relationship between subtest-averaged handshape imagery abilities and the mean RT in the response-to-target predictive task, $r(21) = -0.021$, $p = 0.928$, nonsigners showed a significant relationship between scores, $r(21) = -0.529$, p

= 0.014. Further examination revealed that this correlation was driven by the isolated hand imagery subtest, $r(21) = -0.607$, $p = 0.004$. There was no relationship between nonsigner response-to-target RTs and the functional imagery subtest of the TAMI-h, $r(21) = -0.094$, $p = 0.684$. Thus, nonsigners with better isolated motor imagery abilities exhibited better performance on the response- to-target test.

4. Discussion

The present study was designed to address a number of related questions about the perception of transitional information in the manual modality: (a) To what extent does parsing of manual information rely on transitional phases? (b) During transition periods, are handshape changes informative over and above movement features of the transition? (c) Are sign language users uniquely attuned to handshape change information, in both linguistic and nonlinguistic contexts? And, (d) does motor imagery facilitate the ability to use handshape information to predict an upcoming sign or gesture? We discuss the answers to each of these questions below.

Overall, we found support for a general use of transitional information via faster target detection times for the handshape Blurred video condition that presents only movement cues compared to the Static video condition that provided no transition information at all. We suggest that participants were able to use movement cues alone to narrow the possibilities for the next item. This interpretation is consistent with previous findings using full-signal videos which found that signers could identify signs using transitional cues (e.g. Clark & Grosjean, 1982; Emmorey & Corina, 1990; Grosjean, 1981; ten Holt et al., 2009). Our results further indicate that both signers and nonsigners can use the trajectory information within transitional movements to anticipate an upcoming pseudosign or gesture. Thus, sensitivity to movement trajectory is independent of the linguistic status of the stimulus and linguistic knowledge of the perceiver.

We also found evidence for a signer-specific use of transitional handshape information in the comparison between the Normal video condition in which all transition information is available and the Blurred video condition in which only the handshape information is degraded. For pseudosigns only, signers were better able to detect targets in the Normal than the Blurred condition, while there was no difference between these conditions for the nonsigners for either pseudosigns or grooming gestures. This pattern of results indicates that signers, but not nonsigners, relied on handshape transitions when generating predictions for linguistic stimuli, but not for non-linguistic manual actions. This finding supports the hypothesis that signers use their linguistic knowledge to interpret transitional handshape information during sign recognition and comprehension (e.g., Arendsen et al., 2007; Green, 1984; Hosemann et al., 2013; ten Holt et al., 2009).

We suggest that either (a) signers primarily attend to transitional handshape information when sign phonology is present (as in pseudosigns), or that (b) handshape complexity impacts the utility of transitional information. With respect to the latter hypothesis, we note that the pseudosign stimuli used in this study involved more marked, and therefore more complex, handshapes than their grooming gesture counterparts. Signers may be able to rapidly map complex transitional handshapes to the appropriate linguistic categories via categorical perception (Baker et al., 2005; Emmorey et al., 2003). It may also be the case that the increased handshape complexity present in the pseudosigns resulted in more visually distinct targets for which the signers, but not the nonsigners, had linguistic knowledge and perceptual experience. Additional evidence in favor of this featural complexity explanation can be seen in the post-hoc analysis contrasting the first and second half of trials. Specifically, the nonsigners developed improved use of transitional handshape information for the pseudosigns, but not the grooming gestures.

Signers attended to transitional handshapes for pseudosign stimuli throughout the experiment, with significantly faster response times for Normal compared to Blurred conditions in both the first and second half of the trials. Sign language knowledge may allow individuals to strategically attend to transitional handshape changes from the very beginning of the study, but this attention can be learned without explicit instructions. Nonsigners were able to attend to the same transitional handshape information for pseudosign lists, but only in the latter half of trials. Thus, nonsigners may have gotten better at using handshape complexity in their predictions over time. Future studies might examine whether or not this effect can be bolstered by explicit instruction, akin to Geer and Keane (2017).

In the same split-half analysis, we found that signers also exhibited better performance for the Normal video condition than the Blurred condition for grooming gestures, indicating that they utilized handshape transition information for gestures, but only for the first half of trials. The explanation for this finding is less clear. Sign language experience may prime users to incorporate handshape transitions when making predictions about nonlinguistic manual actions, but doing so without complex handshape information may be taxing. Signers may have adopted strategies over the course of the experiment that made transitional handshapes less informative for grooming gestures.

As noted in the Introduction, we adopted the traditional view of sign language structure in which sign onsets and offsets are defined based on initial/final holds or target locations (e.g., Crasborn et al., 2015; Liddell & Johnson, 1989; Sandler, 1989). The interpretation of our findings would change if we adopted a more linear notion of sign structure in which sign onsets and offsets were defined based on the first and last detectable occurrence of an articulatory feature (Jantunen, 2015). This “longer view” of the sign includes transitions (or parts of them) as part of the lexical sign (or a pseudosign). Given that the pseudosign strings were produced without pauses and with co-articulation, it is likely that there would be few or no videoframes that would be considered purely transitional (i.e., videoframes without any detectable articulatory features that did not belong to the preceding or upcoming pseudosign). Under this longer view of signs, our findings would be re-framed as applying to information contained at the ends and at the beginnings of pseudosigns within a string. In essence, this becomes a definitional problem.

Finally, we utilized the Test of Ability in Movement Imagery for hands (TAMI-h; Donoff et al., 2018) to examine potential effects of motor imagery on transitional information processing, as well as to assess whether sign language experience impacts motor imagery skill. Signers and nonsigners demonstrated similar performance on this measure, and we therefore conclude that sign language experience may not impact motor imagery tasks of this type. Sign language experience may support sensitivity to handshape information in linguistic contexts, but may not affect the ability to imagine isolated hand muscle movements. Future studies might investigate linguistic motor imagery, e.g., for handshapes in one's own vs. another sign language, for example, ASL signers imagining ASL vs. DGS handshapes.

For nonsigners only, there was a significant correlation between scores on the TAMI-h (isolated hand imagery subtest) and target detection RTs for both pseudosigns and grooming gestures. This result indicates that nonsigners, in contrast to signers, relied on motor imagery abilities during target detection. The use of motor imagery may be related to the overall cognitive demand of the response-to-target task for the nonsigners. Signers are accustomed to attending to manual productions. Nonsigners may have needed to devote more attention to monitoring the same productions, incorporating motor imagery into their predictive process.

5. Conclusion

Both signers and nonsigners are able to use transitional movement cues to predict linguistic (pseudosigns) and non-linguistic (grooming gesture) manual actions. However, only signers were able to immediately use transitional handshape information to detect linguistic targets. We hypothesize that signers can use their knowledge of and experience with phonological handshape categories to predict linguistic input. Over time, nonsigners may be able to develop sensitivity to phonological handshapes as their ability to use handshape transitions improved over the course of the experiment. Neither signers nor nonsigners were able to use handshape transition information to predict non-linguistic gesture targets. In addition, signers and nonsigners did not differ in motor imagery abilities, but only nonsigners exhibited a correlation between hand imagery and target prediction skill. We hypothesize that nonsigners may have used motor imagery as a prediction strategy, while signers did not. Overall, our results indicate that signers and nonsigners differ in their ability to incorporate transitional handshape information into predictive representations.

Acknowledgments

This work was supported by a grant from the National Institutes of Health (R01 DC010997). The authors would like express gratitude for the contributions of our participants, our stimulus model, Cindy O'Grady Farnady, Dr. Jörg Matt for statistical consultation, as well as our research assistants: Cami Miner, Israel Montano, and Francynne Pomperada.

Data availability

Data will be made available on request.

References

Arendsen J, van Doorn AJ, & de Ridder H (2007). When and how well do people see the onset of gestures? Gesture, 7(3), 305–342.

Baker SA, Idsardi WJ, Golinkoff RM, & Petitto LA (2005). The perception of handshapes in American sign language. Memory & Cognition, 33(5), 887–904. [PubMed: 16383176]

Brentari D (1998). A prosodic model of sign language phonology. MIT Press.

Brozdowski C, & Emmorey K (2020). Shadowing in the manual modality. Acta Psychologica, 208, Article 103092.

Caselli NK, Sehyr ZS, Cohen-Goldberg AM, & Emmorey K (2017). ASL-LEX: A lexical database of american sign language. Behavior Research Methods, 49(2), 784–801. 10.3758/s13428-016-0742-0 [PubMed: 27193158]

Cheek DA (2001). The phonetics and phonology of handshape in American Sign Language. The University of Texas at Austin.

Clark LE, & Grosjean F (1982). Sign recognition processes in american sign language: The effect of context. Language and Speech, 25(4), 325–340. 10.1177/002383098202500402

Cohen JD (1993). PsyScope: A new graphic interactive environment for designing psychology experiments. Behavioral Research Methods, Instruments, and Computers, 25 (2), 257–271.

Cormier K, Schembri AC, & Tyrone ME (2008). One hand or two?: Nativisation of fingerspelling in ASL and BANZSL. Sign Language & Linguistics, 11(1), 3–44.

Crasborn O, Bank R, Zwitserlood I, Meijer A, Sáfár A, Ormel E, & van der kooij E (2015). Annotation conventions for the Corpus NGT, version 3. 10.13140/RG.2.1.1779.4649

Emmorey K, & Corina D (1990). Lexical recognition in sign language: Effects of phonetic structure and morphology. Perceptual and Motor Skills, 71(3_suppl), 1227–1252. 10.2466/pms.1990.71.3f.1227 [PubMed: 2087376]

Emmorey K, McCullough S, & Brentari D (2003). Categorical perception in American sign language. Language and Cognitive Processes, 18(1), 21–45. 10.1080/01690960143000416

Emmorey K, Midgley KJ, & Holcomb PJ (2022). Tracking the time course of sign recognition using ERP repetition priming. Psychophysiology, 59(3), Article e13975. 10.1111/psyp.13975

Donoff CM, Madan CR, & Singhal A (2018). Handedness effects of imagined fine motor movements. Laterality: Asymmetries of Body, Brain and Cognition, 23(2), 228–248.

Geer L, & Keane J (2014). Exploring factors that contribute to successful fingerspelling comprehension. LREC, 1905–1910.

Geer L, & Keane J (2017). Improving ASL fingerspelling comprehension in L2 learners with explicit phonetic instruction. Language Teaching Research, 1–19.

Green K (1984). Sign boundaries in american sign language. Sign Language Studies, 1042 (1), 65–91. 10.1353/sls.1984.0009

Grosjean F (1981). Sign and word recognition: A first comparison. Sign Language Studies, 32, 195–220. https://www.jstor.org/stable/26203496.

Grosvald M, & Corina DP (2012). The perceptibility of long-distance coarticulation in speech and sign: A study of english and american sign language. Sign Language & Linguistics, 15(1), 73–103.

Johnson RE, & Liddell SK (2011). A segmental framework for representing signs phonetically. Sign Language Studies, 11(3), 408–463. https://www.jstor.org/stable/26190863.

Hosemann J, Herrmann A, Steinbach M, Bornkessel-Schlesewsky I, & Schlesewsky M (2013). Lexical prediction via forward models: N400 evidence from german sign language. Neuropsychologia, 51(11), 2224–2237. [PubMed: 23896445]

IBM Corp. (2017). IBM SPSS statistics for windows (version 25.0). Armonk, NY: IBM Corp. https://hadoop.apache.org.

Jantunen T (2013). Signs and transitions: Do they differ phonetically and does it matter? Sign Language Studies, 13(2), 211–237.

Jantunen T (2015). How long is the sign? Linguistics, 53(1), 93–124.

Keane J, & Brentari D (2016). Fingerspelling: Beyond handshape sequences. In Marschark M, & Spenser P (Eds.), The Oxford handbook of deaf studies in language (pp. 146–160). UK: Oxford University Press.

Kita S, van Gijn I, & van der Hulst H (1998). Movement phases in signs and co-speech gestures, and their transcription by human coders. In Wachsmuth I, & Fröhlich M (Eds.), *Vol.* 1371. Gesture and sign language in human-computer interaction (pp. 23–35). Springer Berlin Heidelberg. 10.1007/BFb0052986.

Klima ES, Tzeng OJL, Fok YYA, Bellugi U, Corina D, & Bettger JG (1999). From sign to script: Effects of linguistic experience on perceptual categorization. Journal of Chinese Linguistics Monograph Series, 13, 96–129. https://www.jstor.org/stable/23826111.

Kutas M, & Federmeier KD (2011). Thirty years and counting: Finding meaning in the N400 component of the event related brain potential (ERP). Annual Review of Psychology, 62, 621.

Liddell SK, & Johnson RE (1989). American Sign Language: The phonological base. Sign Language Studies, 64(1), 195–277.

Mauk CE (2003). Undershoot in two modalities: Evidence from fast speech and fast signing. The University of Texas at Austin.

Ormel E, Crasborn O, & van der Kooij E (2013). Coarticulation of hand height in sign language of the Netherlands is affected by contact type. Journal of Phonetics, 41(3–4), 156–171. 10.1016/j.wocn.2013.01.001

Ormel E, Crasborn O, Kootstra GJ, & de Meier A (2017). Coarticulation of handshape in sign language of the Netherlands: A Corpus study. Laboratory Phonology: Journal of the Association for Laboratory Phonology, 8(1), 1–21. 10.5334/labphon.45, 10.

Palmer SB, Fais L, Golinkoff RM, & Werker JF (2012). Perceptual narrowing of linguistic sign cccurs in the 1st year of life. Child Development, 83(2), 543–553. 10.1111/j.1467-8624.2011.01715.x [PubMed: 22277043]

Pickering MJ, & Garrod S (2013). An integrated theory of language production and comprehension. Behavioral and Brain Sciences, 36(04), 329–347. [PubMed: 23789620]

Recasens D (2018). Coarticulation. In Oxford research encyclopedia of linguistics. 10.1093/acrefore/9780199384655.013.416. Retrieved 13 Jul. 2020.

Sandler W (1989). Phonological representation of the sign: Linearity and nonlinearity in American Sign Language. De Gruyter Mouton.

Schwarz AL (2000). The perceptual relevance of transitions between segments in the fingerspelled signal. University of Texas at Austin. Doctoral dissertation.

StataCorp. (2015). Stata statistical software: Release 14. College Station, TX: StataCorp LP.

Ten Holt GA, Van Doorn AJ, de Ridder H, Reinders MJT, & Hendriks EA (2009). Which fragments of a sign enable its recognition? Sign Language Studies, 9(2), 211–239. 10.1353/sls.0.0012

Wilcox S (1992). The phonetics of fingerspelling. John Benjamins BV.

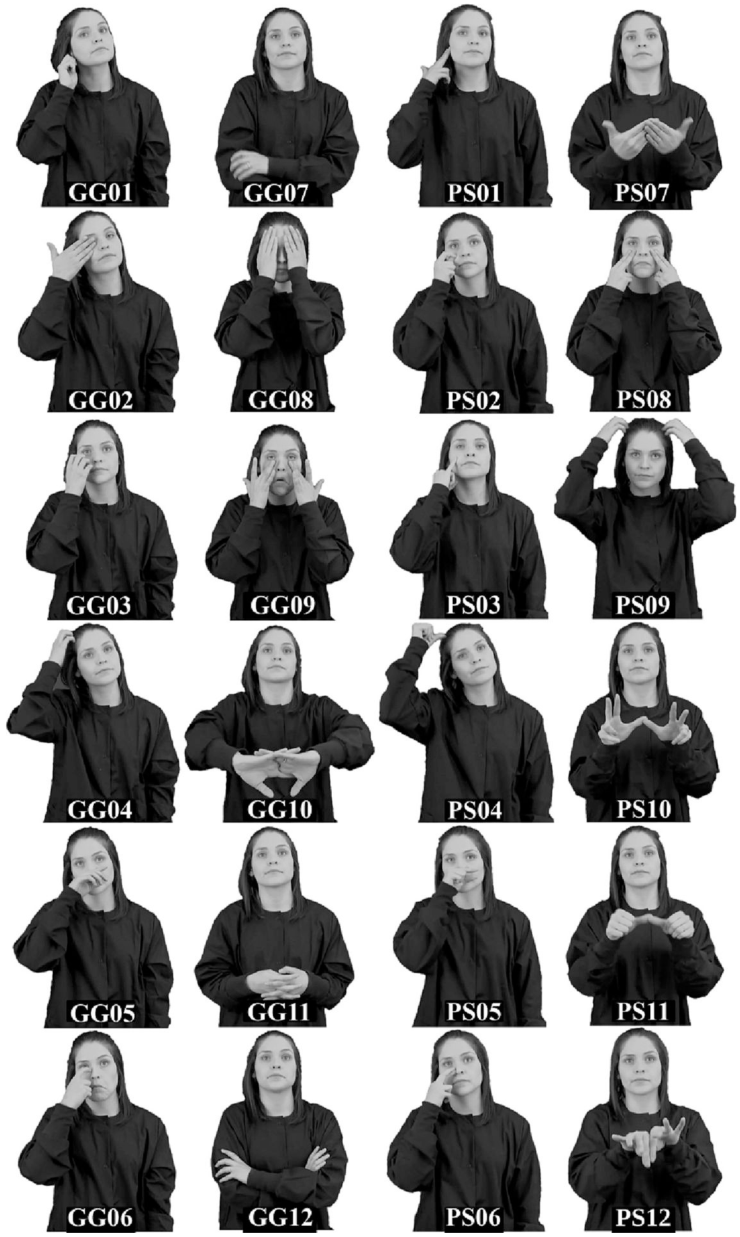

Fig. 1.

Still images of the experimental items for the target-detection task: Grooming Gestures (GG) and Pseudosigns (PS). Full videos are available on OSF: https://osf.io/bzfa5/?view_only=f1ebdd4738a943279c00da5bcaa2c99a.

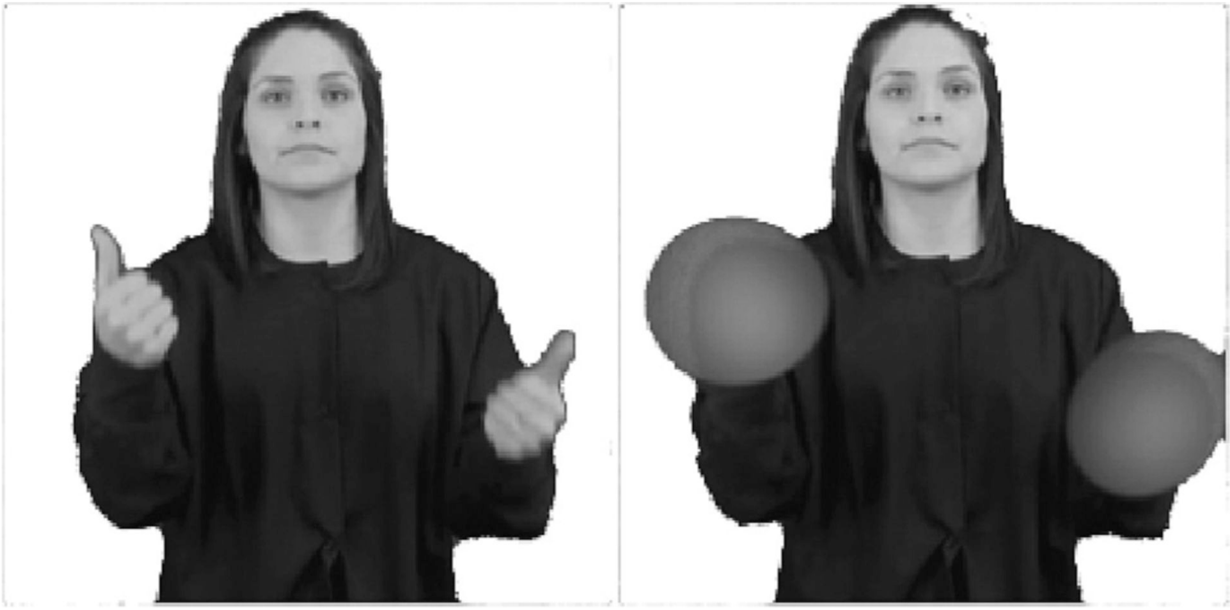

Fig. 2.
Sample video frames from Normal (left) and Blurred (right) pseudosign stimuli. Full videos are available on OSF: https://osf.io/bzfa5/?view_only=f1ebdd4738a943279c00da5bcaa2c99a.

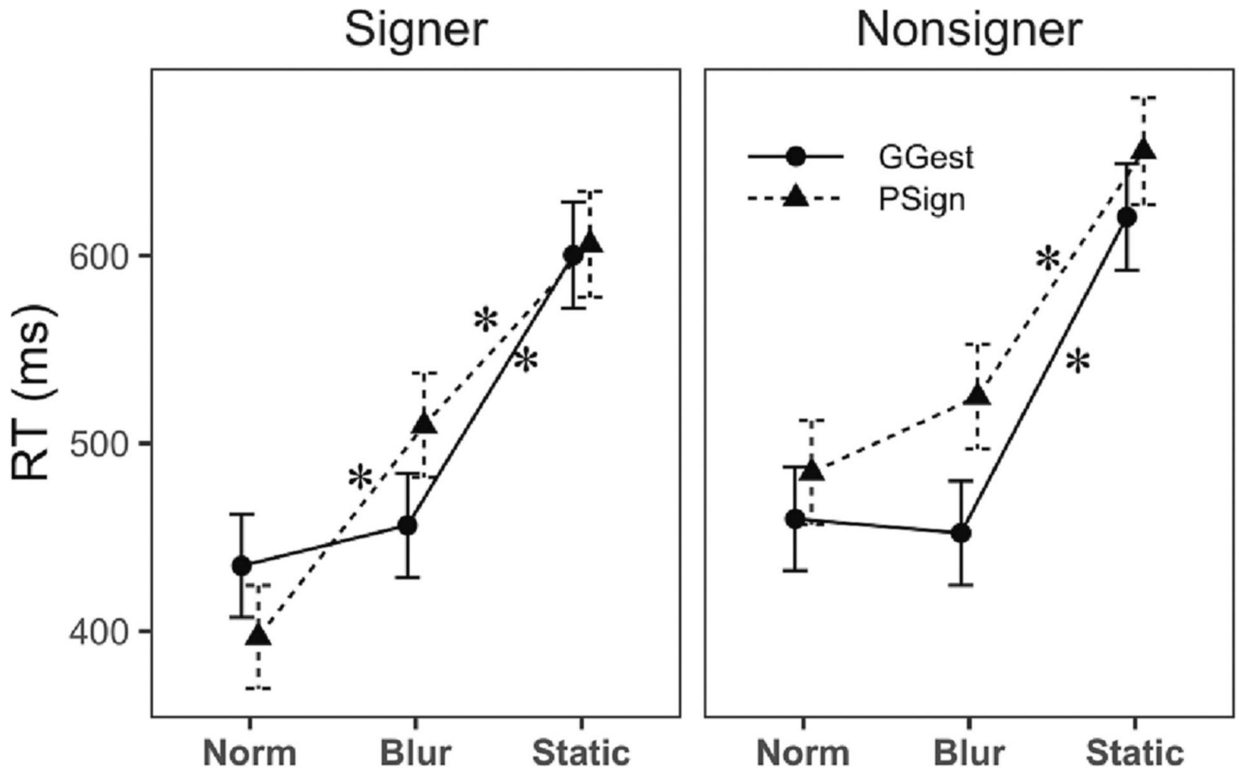

Fig. 3.

Response Time (RT) parameter estimates for a linear mixed effects model examining the effect of Video Condition and Stimulus Type on RTs for each group. Gest = Grooming gesture; PSign = pseudosign; Norm = Normal videos. Stars indicate a significant difference between conditions ($p < 0.05$).

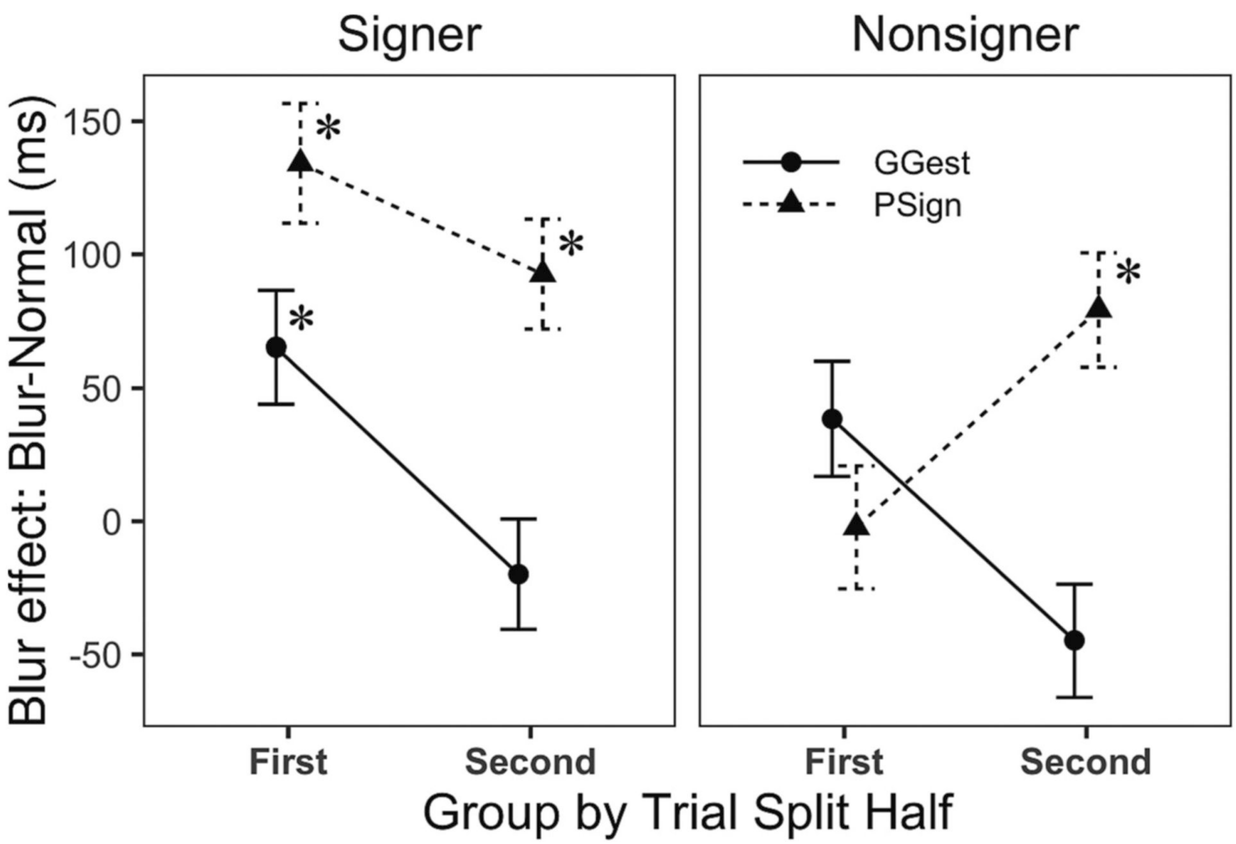

Fig. 4.
Contrasts of marginal linear predictions illustrating the use of transitional handshape information (Blur – Normal conditions) in the first and second halves of the experiment. The linear mixed effects model examined the effect of condition half and type (Gest = grooming gesture; PSign = pseudosign) on RTs for each group. Error bars indicate Standard Error. Stars indicate a significant effect ($p < 0.05$).

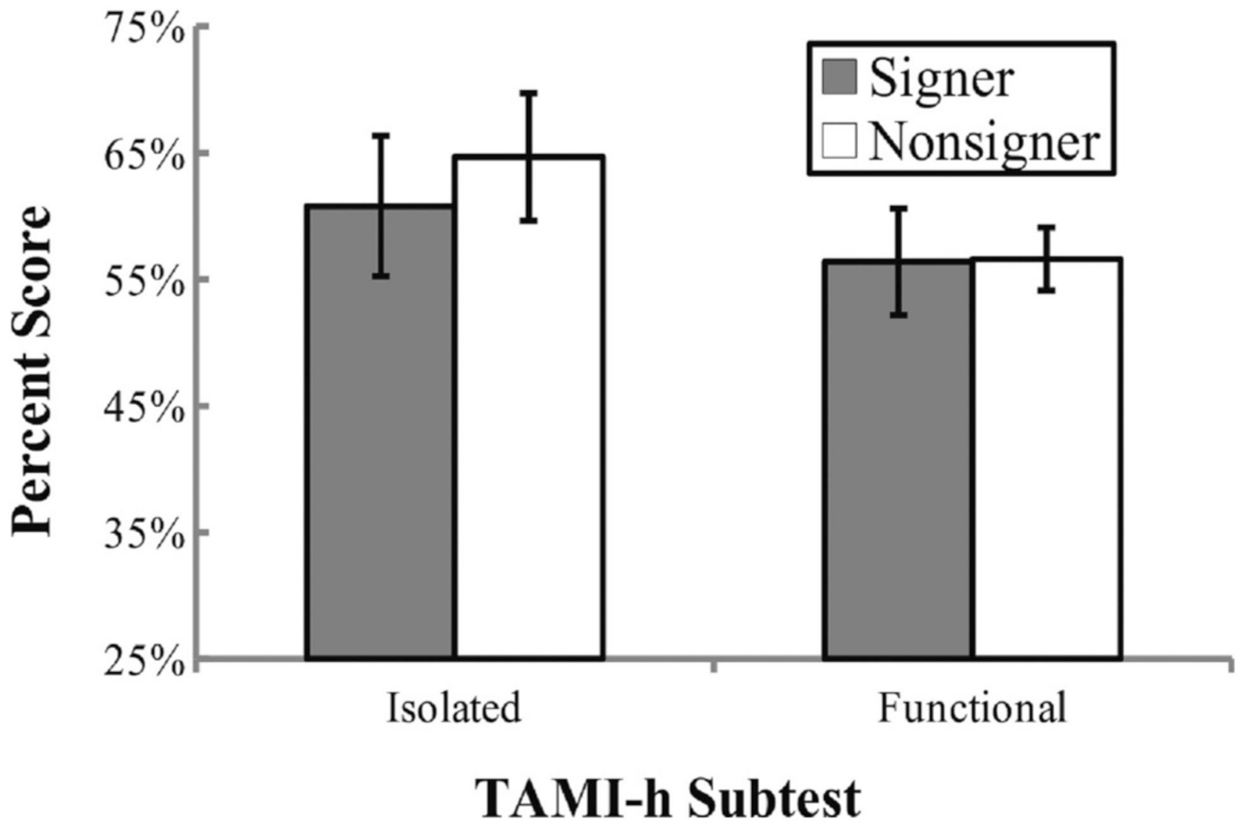

Fig. 5.
Percent correct (mean and standard error) on the hand Test of Ability in Movement Imagery (TAMI-h). For isolated trials, participants match imagined hand configurations with pictures of hands, while for functional imagery trials participants match imagined hand configurations with manipulable objects.

Table 1

LME effect of video condition.

Summary of effect regressions of response times across the fixed factors (model, stimulus type and group), and their interactions.

	Coefficient	SE	z	*p* > |z|
Fixed effects				
Intercept	459.809	27.607	16.66	<0.001
Group	−25.039	38.966	−0.64	0.52
Cond				
Blur - Normal	−7.595	15.366	−0.49	0.621
Static - Normal	160.806	16.333	9.85	<0.001
Group * Cond				
Blur - Normal	29.148	21.546	1.35	0.176
Static - Normal	4.721	22.896	0.21	0.837
Type	24.656	18.967	1.3	0.194
Group * Type	−62.55	26.51	−2.36	0.018
Cond * Type				
Blur - Normal	47.974	21.409	2.24	0.025
Static - Normal	10.266	21.396	0.48	0.631
Group * Cond * Type				
Blur - Normal	43.188	29.901	1.44	0.149
Static - Normal	33.356	29.796	1.12	0.263
Random effects				
Subject				
SD(Intercept)	12,270.34	3085.94		
SD(Type)	0.003	0.019		
SD(Cond)	223.12	189.012		
SD(Item)	17,150.34	1300.198		
SD(Placement)	NA			

Note. SD = Standard Deviation; SE = Standard Error; p = *p*-value significance; Cond = Video condition; Placement = Target Placement within the video; NA = Not available: calculation suppressed due to computational limitations.

Table 2

LME condition split-halves.

Summary of effect regressions of RTs across the fixed factors (model, stimulus type and group), and their interactions, with post-hoc inclusion of split-half condition progress.

	Coefficient	SE	z	*p* > |z|
Fixed effects				
Intercept	425.059	28.772	14.77	<0.001
Group	−16.490	40.584	−0.41	0.685
Cond				
Blur - Normal	38.505	21.631	1.78	0.075
Static - Normal	206.109	22.696	9.08	<0.001
Group * Cond				
Blur - Normal	26.717	30.309	0.88	0.378
Static - Normal	−8.763	31.825	−0.28	0.783
Type	116.970	24.792	4.72	<0.001
Group * Type	−122.647	34.675	−3.54	<0.001
Cond * Type				
Blur - Normal	−40.748	31.461	−1.30	0.195
Static - Normal	−109.423	31.375	−3.49	<0.001
Group * Cond * Type				
Blur - Normal	109.684	43.981	2.49	0.013
Static - Normal	95.698	43.873	2.18	0.029
Half	61.339	23.037	2.66	0.008
Group * Half	−9.848	32.237	−0.31	0.760
Cond * Half				
Blur - Normal	−83.252	30.257	−2.75	0.006
Static - Normal	−79.744	30.722	−2.60	0.009
Group * Cond * half				
Blur - Normal	−1.856	42.417	−0.04	0.965
Static - Normal	18.207	43.042	0.42	0.672
Type * Half	−170.075	30.814	−5.52	<0.001
Group * Type * Half	107.564	42.870	2.51	0.012
Cond * Type * Half				
Blur - Normal	164.707	44.167	3.73	<0.001
Static - Normal	222.021	43.490	5.11	<0.001
Group * Cond * Type * Half				
Blur - Normal	−121.216	61.741	−1.96	0.050
Static - Normal	−112.499	60.645	−1.86	0.064
Random effects				
Intercept	11,295.37	2895.680		
Subject				
SD(Intercept)	1425.50	662.26		

	Coefficient	SE	z	*p* > |z|
SD(Half)	124.67	159.02		
SD(Cond)	< 0.001	0.003		
SD(Type)	16,929.96	1289.71		
Item	11,295.37	2895.68		
Place	NA			

Note. SD = Standard Deviation; SE = Standard Error; p = p-value significance; Cond = Video condition; Half = First or second half of trials; NA = Not available: calculation suppressed due to computational limitations.