Title

Prediction in a visual language: real-time sentence processing in American Sign Language across development

Permalink

https://escholarship.org/uc/item/7fp45641

Journal

Language Cognition and Neuroscience, 33(4)

ISSN

2327-3798

Authors

Lieberman, Amy M
Borovsky, Arielle
Mayberry, Rachel I

Publication Date

2018-04-21

DOI

10.1080/23273798.2017.1411961

Peer reviewed

Prediction in a visual language: real-time sentence processing in American Sign Language across development

Amy M. Lieberman[1], **Arielle Borovsky**[2], and **Rachel I. Mayberry**[3]

[1]Boston University, School of Education, 2 Silber Way, Boston, MA 02215, Phone: (617) 353-3377, alieber@bu.edu

[2]Purdue University, Speech, Language, and Hearing Sciences, 715 Clinic Dr, West Lafayette, IN 47907, Phone: (756) 496-1669, aborovsky@purdue.edu

[3]University of California, San Diego, Department of Linguistics, 9500 Gilman Drive, #0108, La Jolla, CA 92093-0108, Phone: (858) 534-2929, rmayberry@ucsd.edu

Abstract

Prediction during sign language comprehension may enable signers to integrate linguistic and non-linguistic information within the visual modality. In two eyetracking experiments, we investigated American Sign language (ASL) semantic prediction in deaf adults and children (aged 4–8 years). Participants viewed ASL sentences in a visual world paradigm in which the sentence-initial verb was either neutral or constrained relative to the sentence-final target noun. Adults and children made anticipatory looks to the target picture before the onset of the target noun in the constrained condition only, showing evidence for semantic prediction. Crucially, signers alternated gaze between the stimulus sign and the target picture only when the sentential object could be predicted from the verb. Signers therefore engage in prediction by optimizing visual attention between divided linguistic and referential signals. These patterns suggest that prediction is a modality-independent process, and theoretical implications are discussed.

Keywords

Introduction

It is well established that spoken language processing is incremental and dynamic. As listeners perceive and process linguistic input, they actively use the information in the input to anticipate what will come next (see Huettig, Rommers, & Meyer, 2011 for a review). When perceiving a visual scene, listeners can use linguistic input incrementally to narrow down the possible visual referents until sufficient information has been given to identify one unique referent. However, theoretical accounts of how visual and linguistic information is integrated over time are based entirely on a framework where linguistic and non-linguistic information are segregated by sensory modality. That is, spoken language is primarily

perceived via the auditory modality, while the accompanying referential information is typically presented via a visual scene (Tanenhaus, Spivey-Knowlton, Eberhard, & Sedivy, 1995). During the mapping of linguistic input onto a visual scene, linguistic and non-linguistic input are neatly separated by sensory modality, allowing for uninterrupted on-line processing of auditory linguistic and visual nonverbal information simultaneously.

Not all languages allow for such segregation of linguistic and non-linguistic information by sensory modality, however. In contrast to the typical situation where auditory language comprehension and non-linguistic visual information processing is sensorially segregated, individuals who communicate using a sign language such as American Sign Language (ASL) face a different task. ASL is produced manually (using the hands, body, facial expressions and other markers) and perceived visually. To comprehend ASL, signers must focus their visual gaze on the linguistic signal (Emmorey, Thompson, & Colvin, 2008).

The question we investigate here is whether linguistic prediction modulates signers' focus on the linguistic signal relative to the integration of non-linguistic information when both signals occur within the same sensory modality, vision. The demands of on-line, visual sentence processing might preclude the additional uptake of non-linguistic information co-occurring in the visual modality until comprehension is complete. This would suggest that linguistic and non-linguistic information cannot be as readily integrated during on-line comprehension when they occur within the same sensory-modality as compared to when they are segregated by sensory-modality. An alternative possibility, tested here, is that linguistic prediction during on-line comprehension *enables* the integration of linguistic and non-linguistic information when they occur within the same sensory modality. If so, this would suggest that the phenomenon is a fundamental and modality-independent factor in language comprehension.

Investigating gaze patterns during ASL sentence comprehension is crucial to obtain a full theoretical explanation of how prediction affects the integration of linguistic and non-linguistic information. Additionally, identifying signers' gaze patterns during language comprehension can also inform theoretical accounts of how listeners may prioritize visual and linguistic information (Knoeferle & Crocker, 2007) as this provides a unique test case where language and referential processing occurs over the same sensory channel. Before describing the study, we first turn to studies of prediction during language comprehension and the importance of eye gaze in sign language processing in both adults and children.

Prediction during language comprehension

The predictive nature of language processing is demonstrated clearly in experiments where semantic information early in the sentence can be used to predict an upcoming target word at the end of the sentence. Altmann & Kamide (1999) first demonstrated this phenomenon using a visual world eye-tracking paradigm. Participants viewed a scene of an agent (e.g. a boy) and a set of concrete objects (e.g. a cake, a train set, a toy car, and a balloon), and then heard sentences that either constrained the target object based on the verb (e.g. "The boy will eat the cake," in which the cake was the only edible object in the scene), or had no such constraining information (e.g. "The boy will move the cake," in which at least four objects could be moved). Using this paradigm, Altmann & Kamide found that adult participants

were faster to look towards a target in the constrained condition than in the neutral condition. This effect has been replicated in a variety of tasks (Kamide, Altmann, & Haywood, 2003; Nation, Marshall, & Altmann, 2003) and with children as young as age 2 (Fernald, Zangl, Portillo, & Marchman., 2008; Mani & Huettig, 2012), suggesting that predictive processes emerge early in development and may be a central mechanism in language processing (but cf. Huettig & Mani, 2016).

While anticipatory eye-movements and prediction effects during linguistic processing emerge from an early age, there are also important individual differences that appear to influence aspects of prediction ability. Prior studies have found that online predictive abilities correlate with a variety of skills in younger and older children including receptive vocabulary in 3- to 10-year-old children (Borovsky & Creel, 2014; Borovsky, Elman & Fernald, 2012), productive vocabulary in 2-year-olds (Mani & Huettig, 2012), reading ability in 8-year-olds (Mani & Huettig, 2014), literacy skills in adults (Mishra, Singh, Pandey, & Huettig, 2012), and working memory in adults (Huettig & Janse, 2016). Thus, there are individual differences in the degree to which listeners of all ages engage in prediction, and these differences correlate with a variety of abilities tied to linguistic experience and skill.

Gaze and information processing in sign language

In sign language, the task of processing multiple sources of information through vision is acquired early and used routinely to navigate the world. Deaf children learning ASL develop sophisticated strategies for alternating gaze between linguistic information and objects and people in the environment, which enables them to achieve coordinated visual attention with their caregivers (Harris, Clibbens, Chasin & Tibbits, 1989; Waxman & Spencer, 1997). By age two, deaf children with deaf parents show frequent shifts in gaze during mother-child interaction (Lieberman, Hatrak, & Mayberry, 2014). Thus, frequent and meaningful gaze shifts are a natural component of sign language comprehension that develops from an early age.

Despite the modality differences between sign and spoken language, native signers interpret lexical signs much as listeners process spoken words (Bosworth & Emmorey, 2010; MacSweeney et al. 2006; Mayberry, Chen, Witcher, & Klein., 2011). Previous studies of sign lexical processing typically have employed paradigms in which signs are either presented with no visual referents (Emmorey & Corina, 1990; Carreiras, Gutierrez-Sigut, Baquero, & Corina, 2008, Morford & Carlson, 2011), or where the signs and their referents are presented sequentially, such as in priming or picture-matching studies (Bosworth & Emmorey, 2010; Ormel, Hermans, Knoors, & Verhoeven, 2009). To date, there has been little work examining how signers manage visual attention to both a sign stimulus and a concurrent visual scene when the sign stimulus unfolds within a sentence context over time. The current study is a first step in filling this gap.

The current study

Studies of sign language acquisition and processing in typical learners largely suggest modality-independent mechanisms are at play (MacSweeney et al., 2006; Mayberry et al., 2011). Despite these broad similarities, there are also important possible differences in how

signers might integrate visual referents when they must navigate between referents and a linguistic signal within the same modality. While prior studies have explored this question using single words (Lieberman, Borovsky, Hatrak & Mayberry, 2015), they do not address the more fundamental question of how signers engage in on-line sentence comprehension in relation to concurrent non-linguistic information in the visual modality. Given the robustness of the phenomenon of predictive processing in auditory language comprehension, and the consistency with which adults and children have been shown to make anticipatory looks to a target picture in sentences, where the target can be predicted based on semantic information present in the auditory modality, the visual-world paradigm serves as an ideal test case for the current study. Specifically, by contrasting gaze behavior during the comprehension of ASL sentences either with or without a semantically constraining verb, we investigate the role prediction may play on the integration of linguistic and non-linguistic information within the same visual modality.

The visual nature of ASL processing creates competition for visual attention. When the visual system must do "double duty" to recognise information in the both the referential and linguistic streams, it is unknown how signers will direct their visual gaze. If signers apply a strategy of waiting until a discrete unit of linguistic information (i.e. a single sentence) is complete, then they will not shift gaze to a referent until the end of the signed sentence, irrespective of the semantic relation between an earlier verb and later sentential object. This conservative "wait and see" strategy could prove optimal, given recent theoretical proposals that individuals may modulate their language processing strategies based on information available in the moment (Kuperberg & Jaeger, 2016). Alternatively, signers may use prediction the same way as spoken language comprehenders and strategically direct anticipatory looks to the target picture when it can be predicted from the verb (Nation et al., 2003; Mani & Huettig, 2012). This strategy may underlie the visual gaze behavior of deaf children, who have been observed to rapidly shift gaze between linguistic input and visual referents during signed discourse (Lieberman et al., 2014).

Using a stimulus set and paradigm modeled after spoken language semantic prediction studies (Mani & Huettig, 2012), we ask whether semantically constraining information from a verb that predicts an upcoming noun modulates gaze toward a non-linguistic visual target. We compare signers' eye movements while perceiving ASL sentences with a constraining verb such as "EAT[1]" in which the verb EAT constrains the possible target words to one edible object pictured on the screen, to sentences that begin with a sign such as "SEE" in which the verb contains no such constraining information. If visual language processing requires a "wait and see" strategy during on-line sentence comprehension, then signers' gaze patterns to non-linguistic visual targets should not be modulated by the prediction created when the verb semantically constrains the upcoming noun. Alternatively, as we hypothesise, if prediction allows for the integration of linguistic and non-linguistic information within the same modality, as it does when this information is segregated by modality, then signers' gaze patterns should vary as a function of the presence or absence of the semantic constraint created for the upcoming noun from the preceding verb. In Experiment 1, we test this

---

[1]We adopt the convention of using capital letters to represent English glosses for ASL signs.

hypothesis in adult deaf signers. In Experiment 2, we ask whether the gaze patterns observed in adult signers during this sentence processing task are evident in deaf children between the ages of 4 and 8 years old, to determine if there are developmental changes in the timing and consistency of these gaze fixations.

Experiment 1: ASL sentence processing in adult deaf signers

Methods

Participants—Seventeen deaf adults between the ages of 19 to 61 years (M = 32) participated. There were seven females. All of the adults reported using ASL as their primary means of communication. Nine participants had deaf parents and had been exposed to ASL from birth. The remaining eight participants had hearing parents and were first exposed to ASL at various ages—before the age of 2 (n=6), at the age of 5 (n=1) and at the age of 11 (n=1). All participants had been using ASL as their primary form of communication for at least 19 years. One additional adult was tested but was unable to complete the eye-tracking task.

Eye-tracking materials—The stimulus display consisted of four pictures on a screen and one ASL sentence. The stimulus pictures were colorful photo-realistic images presented on a white background square measuring 300 by 300 pixels. The ASL signs were presented on a black background square also measuring 300 by 300 pixels. The pictures and signs were presented on a 17-inch LCD display with a black background, with one picture in each quadrant of the monitor and the sign positioned in the middle of the display, equidistant from the pictures (Figure 1).

Eight sets of four pictures served as the stimuli for the prediction task. Each picture set contained four objects, each of which could be paired with either a neutral verb or a unique semantically-constraining verb. For example, one set consisted of a jug of milk, a baby doll, a book, and a cake, which were paired, respectively, with the verbs POUR, HUG, READ, and EAT. During the experiment, each set of pictures was presented twice—once in the neutral condition and once in the constrained condition, for a total of 16 critical trials. Each picture in the set was equally likely to serve as a target across versions of the stimuli sets. Thus for each participant, one picture from the set of four served as a target in the constrained condition and a different picture served as the target in the neutral condition; each participant saw 16 of the 32 possible target signs. Target items were counterbalanced across participants so that each participant saw eight target nouns produced in the neutral condition and eight different target nouns produced in the constrained condition. The arrangement of pictures was counterbalanced such that each picture was equally likely to appear in any position, and the same picture never occurred in the same location twice. Finally, the order of trials was pseudo-randomised such that the same picture set never appeared in two consecutive trials.

The linguistic stimulus consisted of an ASL sentence that directed the participant towards one of the target pictures. Each sentence was composed of the structure VERB WHAT TARGET. This sentence structure was chosen to be comparable to spoken language sentences typically used in this paradigm such as "See the cake," in which there is a

determiner separating the verb and the noun. In ASL syntax, there is no article before a noun, and use of a pronoun or determiner (e.g. THAT) would necessitate use of a directional cue (i.e. the determiner would be articulated with spatial modification and a non-manual marker). Pragmatically, the WHAT sign following the verb is not intended as a true question or even a rhetorical question, rather it is a syntactic device to focus the constituent (Wilbur, 1996). While adding the sign WHAT to the sentence created a longer verb window, it enabled us to make a clear delineation between the verb and noun windows for analysis.

In the neutral condition, the verb at sentence onset paired equally well with each of the four pictures on the screen. Four different semantically neutral verbs were used -- FIND, LOOK-FOR, SEE, and SHOW-ME. In contrast, in the constrained condition the verb at sentence onset limited the possible target such that only one picture represented an item that was semantically related to the verb. For example, the ASL sentence "POUR WHAT MILK" appeared with only one "pour-able" item (milk) and three unrelated objects that were implausible as objects of the verb POUR (see appendix for a full list of constraining verb-target pairs).

To create the stimulus ASL sentences, a deaf native signer was filmed as she produced each sentence in a natural, child-directed style. The sentences were then edited using Adobe Premiere software to eliminate extraneous frames such that the verb phrase "VERB WHAT" was of uniform length across sentences (2000ms). This allowed us to compare looking time across trials and conditions using a noun onset point of 2000ms following sentence onset. To control for co-articulation effects in the transition from the sign WHAT to the onset of the target word, we chose a conservative definition of sign onset, in which sign onset was defined as the first frame in which the signer's hands left the final position of the sign WHAT. This approach enabled us to account for variation in transition time from the articulation of the previous sign to the initial position of the target sign, such as the difference in time it takes to move the hands to the torso vs. the face. The length of the final noun was 1000ms (+/− 33ms). The video ended on the frame following the final movement of the sign. All sign editing was completed by the first author, a proficient ASL signer, and by a research assistant who is a deaf, native ASL signer.

Experimental task—After obtaining consent, participants were brought into the testing room and seated in front of the LCD display and eye-tracking camera. The stimuli were presented using a PC computer running Eyelink Experiment Builder software (SR Research). Instructions were presented in ASL on a pre-recorded video. Instructions were presented in a child-directed sign register and explained to participants that they would be playing a game, where they would see pictures followed by an ASL sign, and that they should click on the picture that matches the sign. Participants were given two practice trials before the start of the experiment. Next a 5-point calibration and validation sequence was conducted. In addition, a single-point drift check was performed before each trial. The 16 experimental trials were then presented in a single block. After the trial block, participants were given a break during which they watched a short, engaging animated video, and then viewed trials of a different nature as part of a separate experiment.

On each trial, the pictures were first presented on the four quadrants of the monitor. Following a 750ms preview period, a central fixation cross appeared. When the participant fixated gaze on the cross, this triggered the onset of the video sentence stimulus. After the ASL sentence was presented, it disappeared and, following a 500ms interval, a small cursor appeared in the center of the screen. The pictures remained on the screen until the participant clicked on a picture, which ended the trial (Figure 2). The mouse clicking behavior served as confirmation that participants understood the task, but as the cursor did not appear until after the relevant analysis windows, mouse clicking should not influence gaze behavior.

Eye-movement recording—Eye movements were recorded using an Eyelink 1000 remote eye-tracker with remote arm configuration (SR Research) at 500 Hz. The position of the display was adjusted manually such that the display and eye-tracking camera were placed 580–620 mm from the participant's face. Eye movements were tracked automatically using a target sticker affixed to the participant's forehead. Fixations were recorded on each trial beginning at the initial presentation of the picture sets and continuing until the participant clicked on the selected picture. Offline, the data were binned into 50-ms intervals.

Results

Accuracy on the experimental task

All adult participants completed all 16 trials with near 100% accuracy. That is, all participants selected the correct target picture on all 16 trials with the exception of one participant who made one error. This trial was removed from all analyses.

Approach to eye-tracking analysis

The primary goal of the analysis was to determine whether there was an overall effect of prediction, defined as increased looks to the target picture in the constrained condition relative to the neutral condition in the time window that occurs before the onset of the target sign. We then conducted a finer grain inspection of gaze patterns over the time course of the sentence to when and how this prediction effect was manifested.

We divided the time course into two discrete windows of interest for statistical analysis. These windows were derived from the timing of the noun and verb in the linguistic stimulus, as follows. The *verb window* was defined as the portion of the sentence beginning at verb onset and continuing until noun onset. To be consistent with prior spoken language studies (Mani & Huettig, 2012), the analysis began 300ms following verb onset, and extended for 1700ms which was the point of noun onset. During this verb window, any increase in looks to the target in the constrained relative to the neutral condition would be evidence for anticipatory prediction effects. The second window, defined as the *noun window*, began at 300ms following noun onset and continued for 1700ms. This end point corresponded to 4000ms from sentence onset, at which point participants had largely either ended the trial or looked away from the target picture. Figure 2 illustrates the windows of analysis.

To analyze gaze patterns in these discrete time windows, we calculated the proportion of gaze samples in each time window on which participants fixated on the target, and aggregated together all trials within each condition for each participant. To visually examine gaze patterns, we plotted proportion of looks to the target in each window. This measure parallels prior eye-movement analyses in language paradigms using a similar design with children (e.g. Mani & Huettig, 2012), and therefore facilitates comparison more directly across spoken and sign language designs (Figure 3a). However, there are some important methodological differences in our current design compared to spoken visual-world tasks which necessitate a measure that takes into account not only the influence of the target relative to the static pictures, but also accounts for looks to the video of the signer. To address this issue, we transformed the mean proportion of fixations to the target using an empirical logit function (Barr, 2008) calculated by the EyetrackingR package in R (Dink & Ferguson, 2015). The empirical logit allows us to derive a measure that takes into account target fixations with respect to all other interest areas on the screen, including both the dynamic sign video and the static pictures.

We carried out statistical analyses on the empirical logit of target fixations using a linear mixed-effects regression model (Barr, 2008) using the lme4 package in R (Version 3.3.1), with random effects for participants and items (Baayen, Davidson, & Bates, 2008). The model included fixed effects of time window (verb vs. noun, sum-coded and centered), condition (neutral vs. constrained, sum-coded and centered), and the interaction between window and condition. Following the window analysis, we generated a time course visualization of looks to the target, video stimulus, and distractor pictures throughout the 4000ms beginning at sentence onset. Visual inspection of the time course enabled us to observe differences in gaze patterns in the constrained and neutral conditions as the sentence unfolded.

Eye-tracking results

Window analysis: We calculated the mean proportion of looks to the target in the verb and noun time windows in the constrained and neutral conditions (Figure 3). As described above, we fit a linear mixed-effects regression to compare the empirical logit transformation of target fixations by condition (constrained, neutral) and time window (verb, noun). Table 1 outlines the results of this analysis, which revealed main effects of time window and condition, and a significant interaction between time window and condition. Planned pairwise comparisons revealed that participants looked significantly more at the target in the constrained condition versus the neutral condition in the verb window (p = .003) but not in the noun window (p = .6). Thus the data show the expected prediction effect as evidenced by greater (anticipatory) looks to the target during the verb window in the constrained condition only.

Time course of fixations across the sentence: Figure 4 plots the time course of fixations to the target for the first 4000ms following sentence onset. As predicted, the proportion of looks to the target picture increased earlier in the constrained condition than in the neutral condition. Importantly, we also observed an additional gaze pattern unique to sign processing. Specifically, following the initial gaze to the target picture, signers directed a

shift *back* to the linguistic signal, as evidenced by the decline in looks to the target and rise in looks to the ASL sentence between 1000–2000ms following sentence onset. This suggests that signers use a strategy of rapid alternation of gaze between linguistic input and the visual scene to integrate information from both sources of visual input. The robustness of the observed rise and fall in looks to the target during this time window indicates that signers were remarkably consistent in the timing of their gaze shifts from the target picture back to the sentence video. This pattern appears to be an adaptation to dividing attention between visual language and visual referents, a task that is not necessary for listeners perceiving spoken language.

Discussion of Experiment 1—Adult ASL signers showed prediction effects as evidenced by early and sustained looks to the target picture in the semantically constrained condition before the target noun was presented. After the target noun was presented, however, there was no difference in looking time to the target based on condition. Adults clearly used semantically constraining information at the onset of an ASL sentence to predict the target noun occurring at the end of the sentence. Furthermore, adults shifted gaze *back* to the video following an initial gaze to the target picture in the constrained condition. Together these results suggest that deaf adult signers with native or early exposure to ASL rely on prediction as a means to integrate linguistic and non-linguistic information within the same modality, and thus process semantic information incrementally in a way that is highly similar to spoken language processing. Like spoken language processing, ASL processing involves a shift to the target only when prediction enables gaze away from the linguistic signal to the non-linguistic one. Unique to sign language, however, is the fact that this gaze shift *to* the target necessitates a gaze shift *away* from the linguistic signal. Further, following gaze at the target, signers then gaze back at the linguistic stimulus. As discussed below, this may be driven by a desire to ensure that the actual target noun matches the previously predicted target noun, or it may be driven by the visual salience of the dynamic video. Having established that prediction modulates gaze patterns of adult deaf signers during on-line sentence comprehension, we now ask whether the same pattern is observed across development in deaf children.

Experiment 2: ASL sentence processing in child signers

While Experiment 1 demonstrated predictive processing in adult signers, in Experiment 2 we investigated processing in child signers between the ages of 4 and 8. Specifically, we ask whether children in this age range show evidence for predictive processing in the identical sentence comprehension task as stimuli as in Experiment 1. We further ask whether there are individual differences in deaf children's ability to exploit semantically constraining information based on age and/or ASL knowledge. Hearing children processing spoken language make prediction based on semantic information from the age of two; thus we might expect older deaf children to perform in a manner parallel to adults. However, given the increased cognitive demands on visual attention arising from the need to divide attention between linguistic and referential information, children perceiving sign language may not show the same robust anticipatory looking behavior as that exhibited in studies of young spoken language learners.

Methods

Participants

Twenty deaf children (12 F, 8 M) between the ages of 4;2 to 8;1 (M = 6;5, s.d. 1;3) participated. Seventeen children had at least one deaf parent and were exposed to ASL from birth. The remaining three children had two hearing parents and were exposed to ASL by the age of two and a half. All parents reported that they used ASL as the primary form of communication at home. All but two of the children attended a state school for deaf children in which ASL was used as the primary language of instruction; two children attended a segregated classroom for deaf children within a public school program. One additional child was tested but was unable to do the eye-tracking task.

Picture naming task

Each child participant was given a 174-word productive naming task that served as a measure of vocabulary. The task consisted of presenting images to the child on a computer screen and asking the child to produce the ASL sign for each image. The images consisted primarily of concrete nouns (n=163) plus a set of color words (n=11) that represent items used as stimuli in an ongoing series of studies conducted by the authors. The images were chosen to represent common, early-acquired lexical items. Items were shown in groups of 3–6 at a time, and were organised according to semantic categories (i.e. animal signs were shown together; food items were shown together). Participants were not given feedback about accuracy. The task was administered after the eye-tracking task, and was videotaped for later scoring. An item was scored as correct if the participant produced a sign identical to the target sign used in the eye-tracking task.

All 32 pictures that appeared in the current study were included in the picture naming task. This allowed us to verify for each participant whether the child knew the ASL sign for the target items so that we could analyze the data taking into account the child's ability to produce each target sign.[2]

Experimental task

Parental consent was obtained for each child prior to the experiment. The experimental task for the child participants was identical to that of adults, except for the fact that instead of clicking on the matching picture with a mouse, children were instructed to point at the matching picture. This ensured that children did not break gaze from the monitor in order to look at the mouse. A second experimenter sat next to the child to click on the picture to which the child pointed, and to keep the child engaged and encourage them to stay on task.

---

[2]Out of 298 trials in the eye-tracking task, children produced the ASL sign for the target item during picture-naming on 242 trials. The remaining 52 trials contained a target that the child either did not produce or produced with an error during picture-naming. We analyzed the data both with and without these 52 trials, and the pattern of results was identical. Thus, we used the fuller dataset (298 trials) for subsequent analyses of the eye-tracking data.

Results

Picture naming task

The mean performance on the task was 82% (range 32% to 97%). Of the 20 child participants, one participant completed only 117 (out of 174) items due to lack of attention. Importantly, over half of the participants scored at least 90% on the task, suggesting that children were largely familiar with the items that we selected for the study. This is not surprising, as we designed the task to use concrete nouns (and a small number of color words) assumed to be acquired early. Two participants had low scores on the naming task (32% and 33% accuracy). After verifying that the overall pattern of results did not change when these two participants were removed, these participants were not excluded from the subsequent analyses. Nevertheless, the relatively high scores of several participants led us to interpret the picture-naming results as a measure of vocabulary knowledge with caution.

Accuracy on the eye-tracking task

The participants completed all 16 trials except one who completed 12 of the 16 trials. Accuracy ranged from 63% to 100% (corresponding an error rate of 0 to 6 errors), although 14 (out of 20) participants achieved 100% accuracy. Of the 18 incorrect trials, 11 were in the constrained condition, and seven were in the neutral condition. Incorrect trials (5.7% of all trials) were eliminated from all further analyses.

Eye-tracking results

Window analysis—We analyzed empirical logit of looks to the target in the verb and noun windows (Figure 5). We fit a linear mixed-effects regression model to the dataset as described above. Age and vocabulary score (both centered) were additionally entered into the model as predictors to determine whether individual differences contributed to the presence and magnitude of the prediction effect. The dependent variable was the empirical logit transformation of the proportion of fixations to the target in each time window. The fixed effects were condition and time window. Participants and items were entered as random factors. Individual differences were analyzed by entering vocabulary (a centered continuous variable) as a predictor into the model. Parameter estimates indicate a significant effect of time window and condition, and a significant interaction between time window and condition (see Table 2). Planned pairwise comparisons revealed that child participants looked more to the target in the constrained vs. neutral conditions in the verb window (p = .01), but not in the noun window (p = .41). Children showed the expected prediction effect early in the time course, but by the time the noun sign was produced children's looks to the target did not differ by condition. Table 2 shows the results of the regression analysis.

Individual differences—We included age and vocabulary (both centered around their means) as predictor variables in our linear mixed-effects regression model. Vocabulary was defined as the percent score on the picture naming task, i.e. the number of pictures named correctly out of a total of 174 pictures. Neither vocabulary nor age predicted significant variance in the model using empirical logit of target looking. As an exploratory measure, we investigated the effect of vocabulary with looks to target as a proportion of looks to all pictures [T/(T+D)] entered as the dependent measure in the model. Vocabulary was a

significant predictor (Beta estimate = .32, s.e. = .09, t = 3.34, p = .003). Children with higher vocabularies fixated towards the target (relative to the other pictures) more than children with lower vocabularies, although there were no significant interactions between vocabulary and either condition or time window. Age was not a significant predictor in this model. Importantly, vocabulary effects must be interpreted with caution, as two children had very low scores on the task[3] and a high percentage of children were at or near ceiling on this task. This vocabulary effect is discussed further below.

Time course of fixations across the sentence: Figure 6 shows the time course of fixations to the target for the first 4000ms following sentence onset. As predicted, the proportion of looks to the target picture increased earlier in the constrained condition than in the neutral condition. Children also showed alternating gaze between the target and sign stimulus in the constrained condition only.

Experiment 2 Discussion

The gaze patterns of child ASL signers between the ages of 4 and 8 years were modulated by prediction in a largely parallel fashion to the effects of prediction on the gaze patterns of adult signers. Children switched gaze to the target during the preceding verb window only in the constrained condition. After the target noun was signed, there was no difference in target fixations based on condition. Inspection of the time course of gaze fixations suggest that children also alternated gaze between the sign video and target picture in the constrained condition. Importantly, the magnitude of this early predictive gaze shift to the target appears smaller than that in the adult time course. Children's productive vocabulary was positively correlated with target accuracy, but not with empirical logit of target looking. In sum, children showed prediction effects through anticipatory eye movements to the target based on semantically constraining information in the sentence-initial ASL verb.

General Discussion

We tested two hypotheses in two experiments about whether prediction during on-line sentence comprehension modulates the integration of linguistic with non-linguistic information when they occur within the same visual modality. Signers might postpone shifting their gaze until an unfolding ASL sentence is complete in order to avoid the risk of missing portions of the linguistic signal, a risk that does not occur when linguistic and non-linguistic information are segregated by sensory modality as in auditory language comprehension. Alternatively, prediction may enable the integration of linguistic and non-linguistic information within the same modality by reducing the risk of missing linguistic information during gaze shifts away from the linguistic signal to integrate non-linguistic information with sentence comprehension. Prediction would modulate gaze patterns in the latter case but not in the former. Our findings align most closely with the second outcome. We found that deaf adults and children make anticipatory eye movements to a target in a visual world paradigm only when there was semantically constraining information in a sentence-initial verb. Both adult and child signers showed increased looks to a target object

---

[3]When these two participants were removed from the analysis, the effect was in the same direction but was no longer significant.

before the target ASL sign was produced in conditions in which the ASL verb contained semantically constraining information. Adult signers used a clear strategy of alternating gaze between the unfolding linguistic signal and the accompanying referential information in the constrained condition. Child participants showed an emerging pattern resembling that of adults but with greater variability. Finally, we found some evidence that individual differences in child participants predicted the extent to which they were able to exploit semantic information during sentence processing. Overall, these findings demonstrate that prediction arising from semantic constraints is a fundamental property of on-line sentence comprehension. Below, we consider each of the primary findings in detail and their theoretical implications for visually-mediated language processing and the role of prediction in interpreting the linguistic signal.

Prediction and attention during sign language comprehension

Our results clearly demonstrate that ASL signers make anticipatory eye movements to an unnamed object when information in a verb provides sufficient information to identify that target. This was demonstrated through increased looks to the target in the verb time window (i.e. the time window beginning at the onset of the sign verb and ending right before the onset of the noun). As the time course visualization illustrates, the rise in fixations to the target picture occurs only in the constrained condition. In the neutral condition, there are very few looks to any of the pictures as signers continue to watch the video until the target sign is produced. Thus, signers maintain gaze towards the linguistic signal until their prediction is sufficiently strong to allow them to shift gaze toward the non-linguistic information, at which point they begin to shift gaze to the target.

This finding essentially replicates what has been shown in the spoken language literature for adults (Altmann & Kamide, 1999), for typically developing children (Mani & Huettig, 2012), and for children with Specific Language Impairment (SLI; Andreu, Sanz-Torrent, & Trueswell, 2013). To our knowledge, this is the first demonstration that similar predictive processes underlie referential looking behavior during visual, on-line sentence processing in the presence of visual non-linguistic information. Given the robust evidence for the modality-independent nature of language processing, this finding may not seem initially surprising. However, the processing demands of integrating visual non-linguistic information with on-line visual sentence comprehension confronts the signer with a dilemma. Shifting gaze towards a non-linguistic referent requires the comprehender to look away from the linguistic signal – and results in the loss of linguistic information. It might, in fact, be more strategic to delay the re-direction of gaze until the visual sentence is complete. However, this "wait and see" approach is clearly *not* how adult and child signers integrate on-line ASL sentence processing with visual, non-linguistic referents. In fact, signers reliably redirect their gaze away from an ongoing visual linguistic signal only when prediction minimizes the risk that the upcoming linguistic information will alter their sentence comprehension. This suggests that signers, like listeners, are actively and probabilistically balancing the information in the unfolding visual language stream with the information available in the visual world around them (Huettig & Altmann, 2005).

Importantly, the timing of signers' looks away from the linguistic signal generally occurred during the production of the sign WHAT that occurred between the verb and noun. In this way signers may have strategically chosen to look away from the linguistic signal at a time when the signal was minimally informative with regards to identification of the target sign. If this is the case, it would suggest that signers have a sophisticated ability to rapidly assess which source of information (linguistic or non-linguistic) is most informative in the moment, and to time their gaze shifts accordingly. Interestingly, in the neutral condition signers did not glance away from the video during the production of the sign WHAT, perhaps because they did not have sufficient time to look at all four pictures, or because they had no constraining information to narrow down the target. In future work, it will be important to vary the sentence structure— e.g. by having the noun sign immediately follow the verb sign —in order to investigate gaze patterns when looking away from the linguistic signal mid-sentence might cause them to miss meaningful linguistic information.

Linguistic prediction is a strong determining factor underlying decisions about how to allocate attention to language and the visual world. The present findings inform the current debate on the importance of prediction for successful language comprehension (e.g. Huettig & Mani, 2016). In spoken language comprehension studies using eye-movements, listeners can easily gaze towards a target picture without giving up any access to the linguistic signal. It could be argued that individuals make anticipatory looks to a target because there is no competing signal, or no "cost" to making such looks. In fact, listeners will even look at a location where an object was previously viewed even if the object itself is no longer visible (Altmann, 2004). The role of attention has been highlighted in theoretical accounts of listeners' behavior during visual world experiments (Huettig, Mishra, & Olivers, 2011; Kukona & Tabor, 2011; Mayberry, Crocker, & Knoeferle, 2009). In spoken language comprehension, the linguistic signal mediates attention, which is realized behaviorally as eye movements towards a visual referent. In contrast, in the present situation, any anticipatory looks to a picture while the linguistic signal is unfolding results in loss of perception of--or attention to--unfolding linguistic information. Signers only shift gaze away from the visual linguistic signal when semantic constraint reduces this cost. While the findings of the current experiment do not prove that prediction is *necessary* for language comprehension, they demonstrate that at least one of the conditions in which prediction can be selectively deployed is when information available in the linguistic context reduces the cost of diverting attention away from the linguistic signal. Further, the current findings suggest that the attentional mechanisms underlying visual search are robust enough to motivate shifts in gaze from a linguistic signal to a visual referent. Such overt shifting of attention cannot be observed in spoken language paradigms in which attending to the linguistic signal is continuous across the sentence.

Unique patterns in visual language processing

The shift from the ASL sentence to the target, however, does not convey the full picture of how signers perform this task. The second important finding that emerged from the results was a shift from the target picture *back* to the ASL sentence in the constrained condition only. This shift occurred approximately at the onset of the ASL target sign and lasted for only 400–500ms, at which point signers shifted gaze a third time, this time back to the target

picture. This gaze shift was particularly consistent in the adult signers with regard to the timing of the shift relative to the unfolding sentence. This "back-and-forth" pattern contrasts sharply to the one observed in the neutral condition, in which signers watched the ASL sentence without shifting gaze from the onset of the verb until approximately 500ms following the noun onset. This finding indicates that the gaze shifting of signers is yoked to the prediction that occurs during online comprehension. Signers also fixated on the target picture more overall in the constrained condition than in the neutral condition. The time course of fixations suggests that in the constrained condition there was one fixation before the target noun, and a second fixation after the onset of the target noun. Thus the presence of semantic constraints enabled signers to predict the upcoming target, and created a significant difference in the way they allocated visual attention between signs and visual referents.

Shifting gaze is a strategy that signers must use when interacting in environments that include both visual linguistic and visual scene information. For children who are exposed to a natural sign language from birth, this gaze shifting pattern is evident as young as 18 months (Harris et al., 1989), and by the age of two to three children rapidly and purposefully shift gaze every few seconds during a parent-child book reading interactions (Lieberman et al., 2014). By the time they are school age, and certainly by the time they are adults, this strategy of gaze shifting is one they have consistently applied to visual language processing. Nevertheless, in the current study, there were several alternative possibilities for how signers may have approached the task. Signers could have deduced from the first trials of the study that when the verb contained semantically constraining information, the noun that followed at the end was always the one that was expected (i.e. all sentences were very high cloze probability, and there were no instances of semantic violation where the sentence-final noun did not match the earlier verb). Instead, signers adopted a clear strategy of shifting gaze during sentences where the verb constrained the target identity, and timed that gaze shift to coordinate with production of the sign WHAT which reliably occurred between the verb and the noun. In sentences where the target could not be determined until the end of the sentence, and prediction could not occur, signers employed a different strategy; they fixated on the sign sentence until the target word was produced.

It is important to note that reliance on visual linguistic information during language processing is not completely absent from spoken language comprehension. In typical face-to-face interaction listeners rely on visual information from the face to interpret the linguistic signal. Even pre-lingual infants are sensitive to audio-visual synchrony in perceiving speech sounds (Kuhl & Meltzoff, 1984; Weikum, Vouloumanos, Navarra, Soto-Faraco, Sebastián-Gallés, & Werker, 2007). When linguistic information is presented multi-modally, children reliably use speaker gaze as an informative cue to interpret a novel verb, particularly when the accompanying linguistic signal is referentially uninformative (Nappa, Wessel, McEldoon, Gleitman, & Trueswell, 2009). However, the amount of visual information on which participants must rely depends on the corresponding informativeness of the auditory cue when processing spoken language. In contrast, when processing sign language, individuals must obtain all linguistic information visually, and thus they exhibit the observed unique pattern of predictive gaze-shifting. In future work, it will be important to evaluate how non-signers process multi-modal information in a similar paradigm, for example using audio-visual spoken language input, gestures, or written words as the

linguistic stimulus. However, we speculate that the shifting patterns exhibited here as a result of prediction are representative of the way that all signers navigate the real world in which linguistic and visual information are consistently perceived simultaneously through only one sensory channel - vision.

Language development and gaze

Finally, the child participants in our study showed effects of vocabulary in their ability to use semantically constraining information to generate predictions allowing them to make anticipatory eye movements to a target. This effect emerged when evaluating accuracy in looking at the target relative to the distractors, but may have been at least partially driven by a small number of participants who had a low score on the picture naming task. We speculate that vocabulary would have emerged as a stronger predictor if our language measure reflected the true range of vocabulary abilities in our sample (Rommers, Meyer & Huettig, 2015). The lack of established vocabulary measures for deaf children older than 36 months limited our ability to obtain a more sensitive measure of children's vocabulary (Allen & Enns, 2013; Anderson & Reilly, 2002).

Given the wide age range of the child participants, it was initially somewhat surprising that age was not a significant predictor of performance on the sentence processing task. However, when considering the present results against previous findings with children in a similar age range (Borovsky, et al., 2012), it is likely that children in the age range tested had acquired the ability to integrate semantic information in order to make anticipatory eye movements towards a target. In fact, the current findings align with previous findings that the mechanisms of prediction are largely in place in children as young as two (Mani & Huettig, 2012). Finally, the sentences in the current study were syntactically simple and comprised only three signs, suggesting that looking behavior, comprehension, and prediction during ASL sentence processing are all in place when basic vocabulary and sentence structure have been acquired at an early age. A productive avenue for future research would be to explore how prediction in more grammatically complex instances changes with development.

In the current study, all children had been exposed to ASL either from birth or from early in life, which does not represent the experience of the majority of deaf children born today (Mitchell & Karchmer, 2004). Thus the correlation with vocabulary is particularly significant given that many deaf children face substantial delay in exposure to a first language. Beyond the processing of single sentences in the current task, deaf children must become adept at making the moment-to-moment decisions about where to look and when to look during interaction with people and the world. Clearly the ability to integrate visual information from many sources is critical for deaf children. As the present results show, prediction based on on-line sentence comprehension plays a major role in their ability to accomplish this within-modality integration of linguistic and non-linguistic information. A remaining empirical question is whether this looking behavior can only be learned when sign language is acquired early, or whether it is a simple product of learning to comprehend sign language, regardless of age. If a lack of experience with sign language leads to a

decreased ability to allocate visual attention, this could have important implications for deaf children's ability to manage in the visually complex environments they encounter.

Conclusion

The current study demonstrated that deaf adults and children show predictive processing during real-time comprehension of ASL sentences, and make anticipatory eye movements to a target based on semantic information provided early in the sentence. Signers integrated visual linguistic information and referential scene information through rapid shifts between the unfolding linguistic signal and the corresponding target that were modulated by prediction generated during sentence comprehension. Adults showed a robust pattern of gaze alternation throughout the sentence; children showed a similar pattern in emerging form. Children's vocabulary level was a significant predictor of semantic processing ability. These findings demonstrate that when linguistic and referential cues exist in the same modality, signers' gaze patterns are driven by their ability to predict the referent of an unfolding sentence. This pattern indicates that sign language incorporates similar predictive mechanisms of language processing as in spoken language interpretation, and serves as a first step in understanding how perception may be allocated between language and the visual world when both are perceived through a single modality. Importantly, our results also contribute to theoretical accounts of prediction in language, which have, until now, been informed primarily from research in the spoken domain. While additional work is necessary, our work represents an initial, important step in outlining the degree to which prediction may be a fundamental mechanism in guiding attention during language processing, irrespective of modality.

References

Allen TE, Enns C. A psychometric study of the ASL Receptive Skills Test when administered to deaf 3-, 4-, and 5-year-old children. Sign Language Studies. 2013; 14(1):58–79. DOI: 10.1353/sls. 2013.0027

Altmann GT. Language-mediated eye movements in the absence of a visual world: The 'blank screen paradigm'. Cognition. 2004; 93(2):B79–B87. DOI: 10.1016/j.cognition.2004.02.005 [PubMed: 15147941]

Altmann GT, Kamide Y. Incremental interpretation at verbs: Restricting the domain of subsequent reference. Cognition. 1999; 73(3):247–264. DOI: 10.1016/S0010-0277(99)00059-1 [PubMed: 10585516]

Anderson D, Reilly J. The MacArthur communicative development inventory: normative data for American Sign Language. Journal of Deaf Studies and Deaf Education. 2002; 7(2):83–106. DOI: 10.1093/deafed/7.2.83 [PubMed: 15451878]

Andreu L, Sanz-Torrent M, Trueswell JC. Anticipatory sentence processing in children with specific language impairment: Evidence from eye movements during listening. Applied Psycholinguistics. 2013; 34(01):5–44. DOI: 10.1017/S0142716411000592

Baayen RH, Davidson DJ, Bates DM. Mixed-effects modeling with crossed random effects for subjects and items. Journal of Memory and Language. 2008; 59(4):390–412. DOI: 10.1016/j.jml. 2007.12.005

Barr DJ. Analyzing 'visual world' eyetracking data using multilevel logistic regression. Journal of Memory and Language. 2008; 59(4):457–474. DOI: 10.1016/j.jml.2007.09.002

Borovsky A, Creel SC. Children and adults integrate talker and verb information in online processing. Developmental Psychology. 2014; 50(5):1600–1613. DOI: 10.1037/a0035591 [PubMed: 24611671]

Borovsky A, Elman JL, Fernald A. Knowing a lot for one's age: Vocabulary skill and not age is associated with anticipatory incremental sentence interpretation in children and adults. Journal of Experimental Child Psychology. 2012; 112(4):417–436. DOI: 10.1016/j.jecp.2012.01.005 [PubMed: 22632758]

Bosworth RG, Emmorey K. Effects of iconicity and semantic relatedness on lexical access in American Sign Language. Journal of Experimental Psychology: Learning, Memory, and Cognition. 2010; 36(6):1573.doi: 10.1037/a0020934

Carreiras M, Gutiérrez-Sigut E, Baquero S, Corina D. Lexical processing in Spanish sign language (LSE). Journal of Memory and Language. 2008; 58(1):100–122. DOI: 10.1016/j.jml.2007.05.004

Dink, JW., Ferguson, B. eyetrackingR: An R Library for Eye-tracking Data Analysis. 2015. Retrieved from http://www.eyetrackingr.com

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

Emmorey K, Thompson R, Colvin R. Eye gaze during comprehension of American Sign Language by native and beginning signers. Journal of Deaf Studies and Deaf Education. 2008; 14(2):237–243. DOI: 10.1093/deafed/enn037 [PubMed: 18832075]

Fernald, A., Zangl, R., Portillo, AL., Marchman, VA. Looking while listening: Using eye movements to monitor spoken language. In: Sekerina, IA.Fernández, EM., Clahsen, H., editors. Developmental Psycholinguistics: On-line Methods in Children's Language Processing. Vol. 44. John Benjamins Publishing; 2008. p. 113-132.

Harris M, Clibbens J, Chasin J, Tibbitts R. The social context of early sign language development. First Language. 1989; 9(25):81–97. DOI: 10.1177/014272378900902507

Huettig F, Altmann GT. Word meaning and the control of eye fixation: Semantic competitor effects and the visual world paradigm. Cognition. 2005; 96(1):B23–B32. DOI: 10.1016/j.cognition. 2004.10.003 [PubMed: 15833303]

Huettig F, Janse E. Individual differences in working memory and processing speed predict anticipatory spoken language processing in the visual world. Language, Cognition and Neuroscience. 2016; 31(1):80–93. DOI: 10.1080/23273798.2015.1047459

Huettig F, Mani N. Is prediction necessary to understand language? Probably not. Language, Cognition and Neuroscience. 2016; 31(1):19–31. DOI: 10.1080/23273798.2015.1072223

Huettig F, Rommers J, Meyer AS. Using the visual world paradigm to study language processing: A review and critical evaluation. Acta psychologica. 2011; 137(2):151–171. DOI: 10.1016/j.actpsy. 2010.11.003 [PubMed: 21288498]

Kamide Y, Altmann GT, Haywood SL. The time-course of prediction in incremental sentence processing: Evidence from anticipatory eye movements. Journal of Memory and language. 2003; 49(1):133–156. DOI: 10.1016/S0749-596X(03)00023-8

Knoeferle P, Crocker MW. The influence of recent scene events on spoken comprehension: Evidence from eye movements. Journal of Memory and Language. 2007; 57(4):519–543. DOI: 10.1016/j.jml.2007.01.003

Kukona A, Tabor W. Impulse processing: A dynamical systems model of incremental eye movements in the visual world paradigm. Cognitive science. 2011; 35(6):1009–1051. DOI: 10.1111/j. 1551-6709.2011.01180.x [PubMed: 21609355]

Kuperberg GR, Jaeger TF. What do we mean by prediction in language comprehension? Language, cognition and neuroscience. 2016; 31(1):32–59. DOI: 10.1080/23273798.2015.1102299

Kuhl PK, Meltzoff AN. The intermodal representation of speech in infants. Infant Behavior and Development. 1984; 7(3):361–381. DOI: 10.1016/S0163-6383(84)80050-8

Lieberman AM, Borovsky A, Hatrak M, Mayberry RI. Real-time processing of ASL signs: Delayed first language acquisition affects organization of the mental lexicon. Journal of Experimental Psychology: Learning, Memory, and Cognition. 2015; 41(4):1130–1139. DOI: 10.1037/xlm0000088

Lieberman AM, Hatrak M, Mayberry RI. Learning to look for language: Development of joint attention in young deaf children. Language Learning and Development. 2014; 10(1):19–35. DOI: 10.1080/15475441.2012.760381

MacSweeney M, Campbell R, Woll B, Brammer MJ, Giampietro V, David AS, Calvert GA, McGuire PK. Lexical and sentential processing in British Sign Language. Human Brain Mapping. 2006; 27(1):63–76. DOI: 10.1002/hbm.20167 [PubMed: 15966001]

Mani N, Huettig F. Prediction during language processing is a piece of cake—But only for skilled producers. Journal of Experimental Psychology: Human Perception and Performance. 2012; 38(4): 843.doi: 10.1037/a0029284 [PubMed: 22774799]

Mani N, Huettig F. Word reading skill predicts anticipation of upcoming spoken language input: A study of children developing proficiency in reading. Journal of Experimental Child Psychology. 2014; 126:264–279. DOI: 10.1016/j.jecp.2014.05.004 [PubMed: 24955519]

Mayberry MR, Crocker MW, Knoeferle P. Learning to attend: A connectionist model of situated language comprehension. Cognitive science. 2009; 33(3):449–496. DOI: 10.1111/j. 1551-6709.2009.01019.x [PubMed: 21585477]

Mayberry RI, Chen JK, Witcher P, Klein D. Age of acquisition effects on the functional organization of language in the adult brain. Brain and Language. 2011; 119(1):16–29. DOI: 10.1016/j.bandl. 2011.05.007 [PubMed: 21705060]

Mishra RK, Singh N, Pandey A, Huettig F. Spoken language-mediated anticipatory eye-movements are modulated by reading ability-Evidence from Indian low and high literates. Journal of Eye Movement Research. 2012; 5(1)doi: 10.16910/jemr.5.1.3

Mitchell RE, Karchmer MA. Chasing the mythical ten percent: Parental hearing status of deaf and hard of hearing students in the United States. Sign Language Studies. 2004; 4(2):138–163. DOI: 10.1353/sls.2004.0005

Morford JP, Carlson ML. Sign perception and recognition in non-native signers of ASL. Language learning and development. 2011; 7(2):149–168. DOI: 10.1080/15475441.2011.543393 [PubMed: 21686080]

Nappa R, Wessel A, McEldoon KL, Gleitman LR, Trueswell JC. Use of speaker's gaze and syntax in verb learning. Language Learning and Development. 2009; 5(4):203–234. DOI: 10.1080/15475440903167528 [PubMed: 24465183]

Nation K, Marshall CM, Altmann GT. Investigating individual differences in children's real-time sentence comprehension using language-mediated eye movements. Journal of Experimental Child Psychology. 2003; 86(4):314–329. DOI: 10.1016/j.jecp.2003.09.001 [PubMed: 14623215]

Ormel E, Hermans D, Knoors H, Verhoeven L. The Role of Sign Phonology and Iconicity during Sign Processing: The Case of Deaf Children. Journal of Deaf Studies and Deaf Education. 2009; 14(4): 436–448. DOI: 10.1093/deafed/enp021 [PubMed: 19620251]

Rommers J, Meyer AS, Huettig F. Verbal and nonverbal predictors of language-mediated anticipatory eye movements. Attention, Perception, & Psychophysics. 2015; 77(3):720–730. DOI: 10.3758/ s13414-015-0873-x

Tanenhaus MK, Spivey-Knowlton MJ, Eberhard KM, Sedivy JC. Integration of visual and linguistic information in spoken language comprehension. Science. 1995; 268(5217):1632.doi: 10.1126/ science.7777863 [PubMed: 7777863]

Waxman RP, Spencer PE. What mothers do to support infant visual attention: Sensitivities to age and hearing status. Journal of Deaf Studies and Deaf Education. 1997; 2(2):104–114. DOI: 10.1093/ oxfordjournals.deafed.a014311 [PubMed: 15579840]

Weikum WM, Vouloumanos A, Navarra J, Soto-Faraco S, Sebastián-Gallés N, Werker JF. Visual language discrimination in infancy. Science. 2007; 316(5828):1159–1159. DOI: 10.1126/science. 1137686 [PubMed: 17525331]

Wilbur, R. Evidence for the function and structure of wh-clefts in American Sign Language. In: Edmondson, W., Wilbur, RB., editors. International review of sign linguistics. Vol. 1. Mahwah, NJ: Lawrence Erlbaum Associates, Inc.; 1996. p. 209-256.
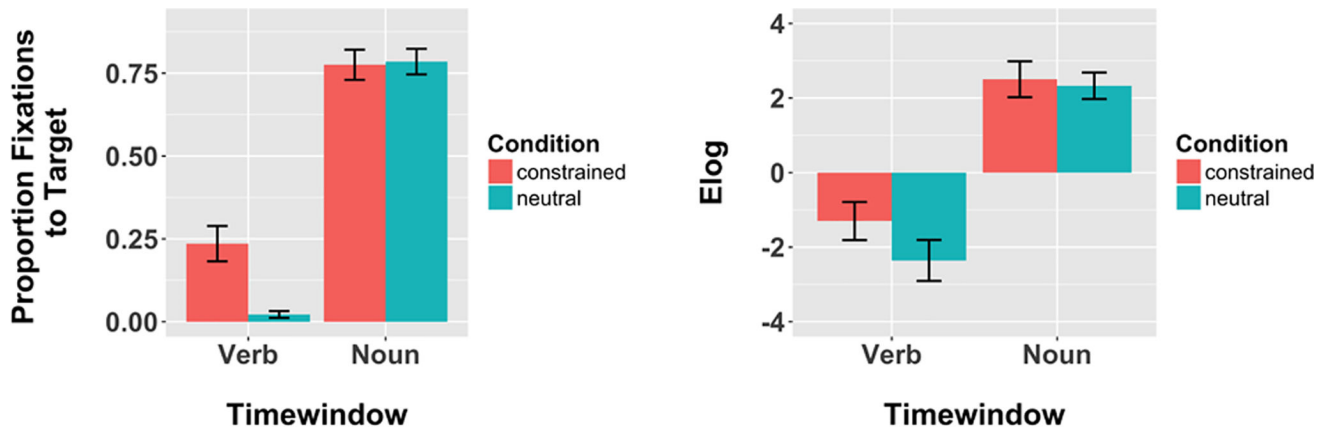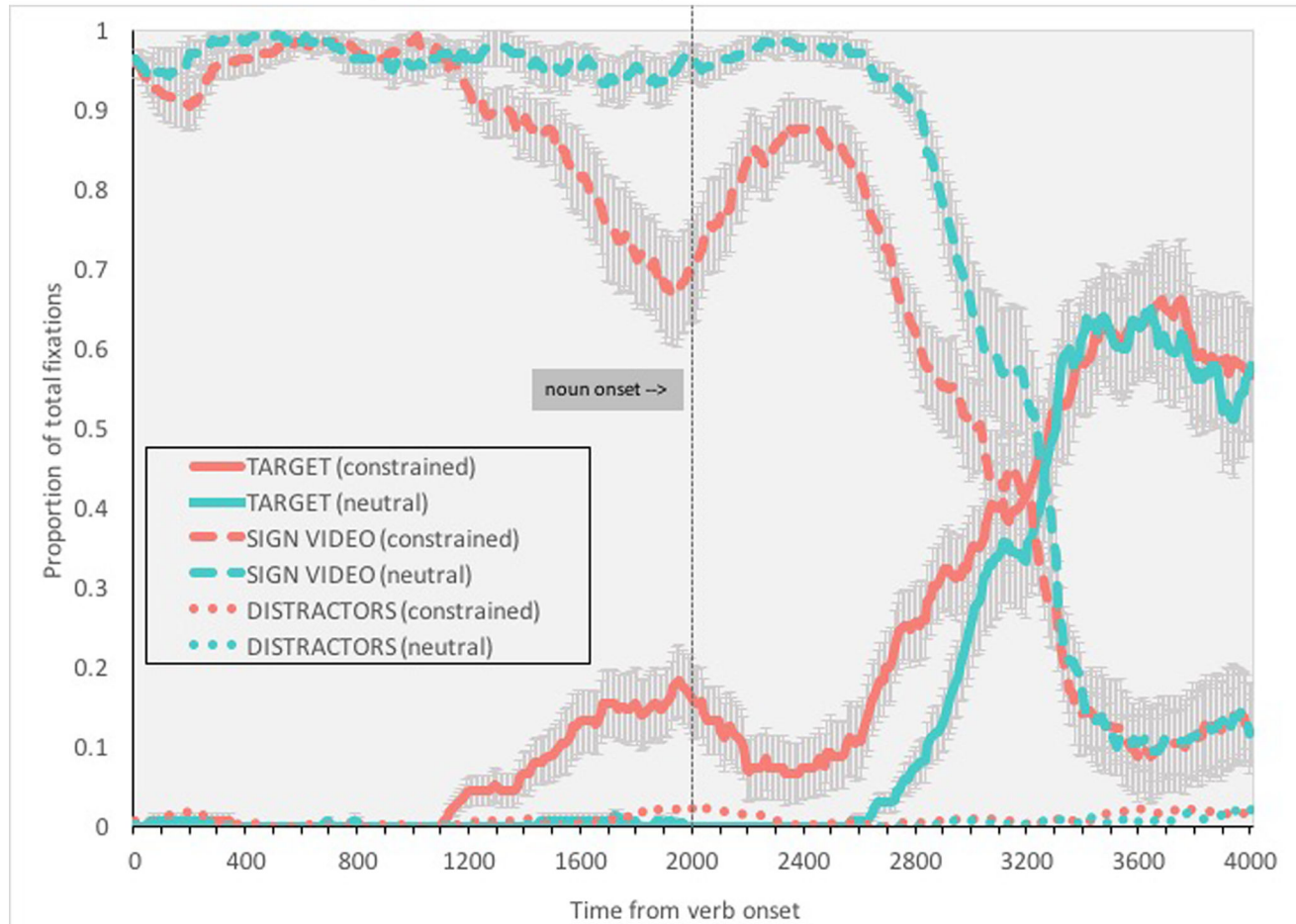
Figure 1.
Layout of pictures and signed video stimuli
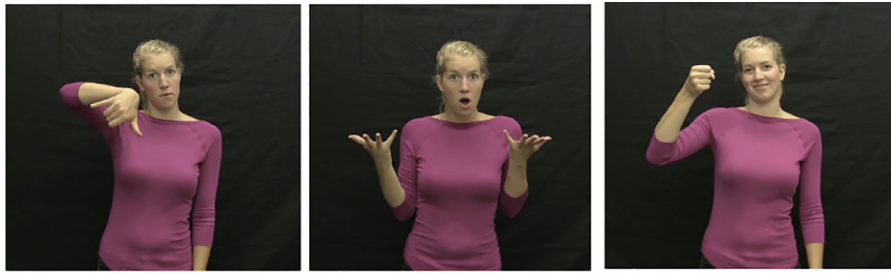
Figure 2.
Schematic of time periods used for statistical analysis

Figure 3.
Adults' mean fixations (s.e.) to target by time window and condition; a) proportion and b) empirical logit (Elog) of target looks.
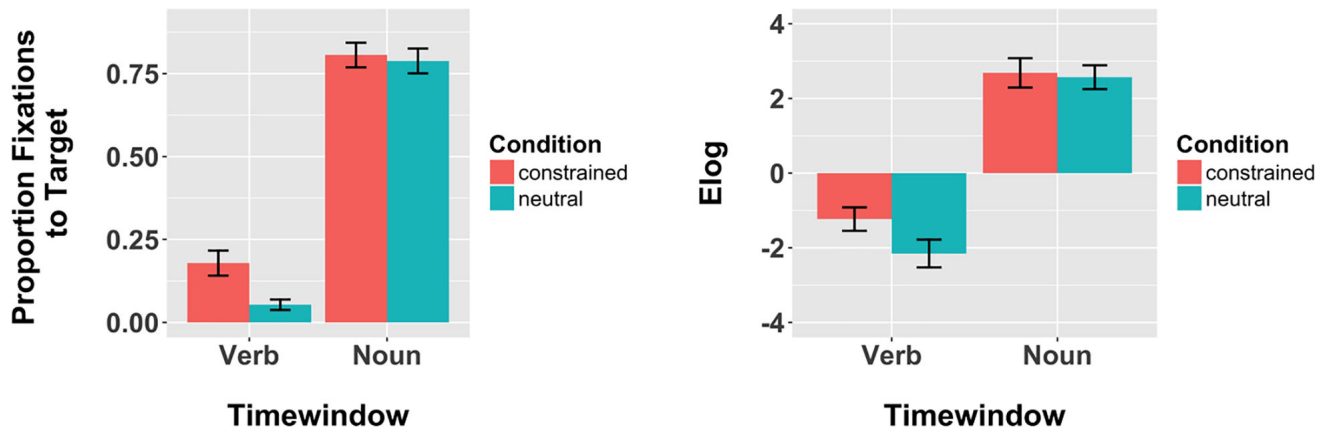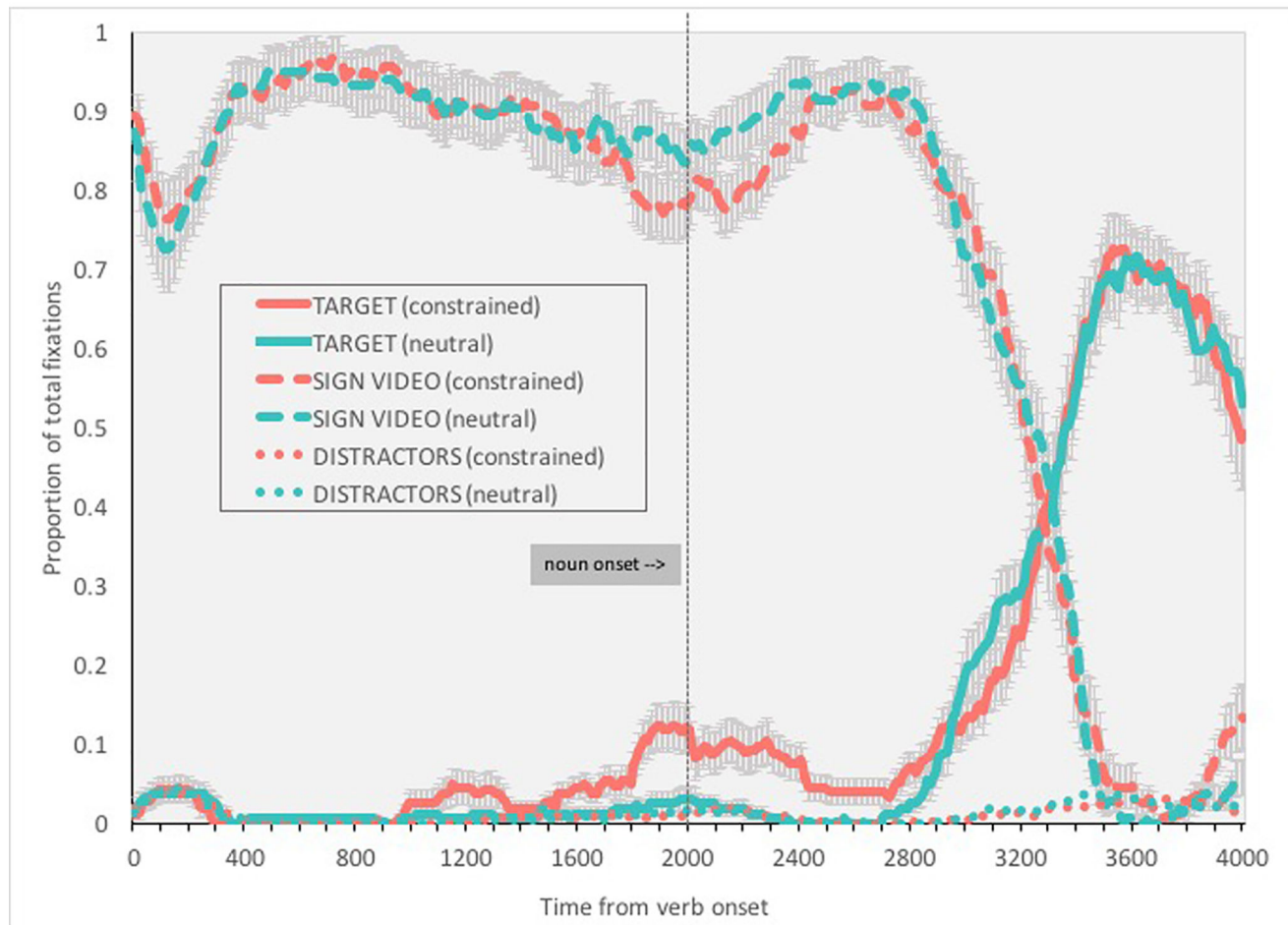
Figure 4.
Time course of adults' mean fixations to the sign video, target picture, and distractor pictures from 0-4000ms following verb onset. The vertical line at 2000ms marks the end of the sign WHAT and the onset of the target noun.

Figure 5.
Children's mean fixations (s.e.) to target by time window and condition; a) accuracy and b) empirical logit of target looks.

POUR WHAT?   MILK.

Figure 6.
Time course of children's mean fixations to the sign, target picture, and distractor pictures from 0-4000ms following verb onset. The vertical line at 2000ms marks the end of the sign WHAT and the onset of the target noun.

Table 1

Parameter estimates from the best-fitting mixed-effects regression model of the effects of time window and condition on empirical logit of adults' looks to the target picture

Fixed effects	Estimate	Std. Error	t value	p-value
(Intercept)	3.40	.20	17.12	<.001
Condition	−2.23	.31	−7.08	<.001
Time window	−2.74	.26	−10.56	<.001
Condition x Time window	−3.17	.52	−6.13	<.001

Table 2

Parameter estimates from the best-fitting mixed-effects regression model of the effects of time window, condition, age and vocabulary on elog of children's looks to the target picture

Term	Estimate	Std. Error	t value	p value
(Intercept)	3.65	.13	28.61	<.001
Condition	−.55	.18	−3.13	.002
Time window	−1.82	.16	−11.52	<.001
Cond. X Time Window	−.91	.31	−2.89	.004
Vocabulary score	−.05	.68	−.08	.94
Age	.01	.01	1.28	.22