Real-time lexical comprehension in young children learning American Sign Language (ASL)

Kyle MacDonald

Todd LaMarr

David Corina

Virginia A. Marchman

Anne Fernald

**Word count: 2110** = acknowledgements, +introduction, + discussion

# Author note:

We are grateful to the children and parents who participated in this research, and to the California School for the Deaf in Fremont. Special thanks to Karina Pedersen, Lisalee Egbert , Laura Petersen, and Michele Berke for their help with participant recruitment; to Pearlene Utley and Rosa Lee Timm for help with creation of the stimuli; to Michael Henry Tessler for help with data analysis; and to Shane Blau, Kat Adams, Melanie Ashland, and the staff of the Language Learning Lab at Stanford University for helpful discussions. This work was made possible by an NIDCD grant to Anne Fernald and David Corina (R21 DC012505).

**Abstract**

The ability to interpret language rapidly from moment to moment as the signal unfolds in time is critical for developing language proficiency. Research on spoken language learning by very young children has used eye movements during real-time sentence processing as a window into their emerging comprehension abilities (Fernald & Marchman, 2012). But, we know relatively little about how children learning signed languages develop proficiency in real-time visual language comprehension. In this study, we developed the first measures of young children’s abilities in real-time comprehension of American Sign Language (ASL), investigating how these language processing skills are linked to children’s age, vocabulary, and hearing status. Participants were 29 native ASL-learning children (16-53 mos, 16 deaf and 13 hearing) and fluent adult signers (*n*=19). When assessed in a task of real-time ASL comprehension, children’s comprehension skills improved with age, moving toward the efficiency of adult signers. Importantly, children’s processing skills were associated with vocabulary size, showing that the ability to establish reference in real time is linked to language learning. Finally, we found that both deaf and hearing ASL learners showed qualitatively similar patterns of looking behavior, suggesting that their visual language processing skills are driven by experience with a visual language, and not by deafness. These findings show parallels between children learning signed and spoken languages in the early development of proficiency in real-time language comprehension.

**Real-time lexical comprehension in young children learning**

**American Sign Language (ASL)**

Learning to find meaning in a spoken or a signed language requires learning to establish reference during real-time interaction – relying on audition to interpret spoken words, and on vision to interpret manual signs. Starting in infancy, children learning a spoken language make dramatic gains in their ability to link acoustic signals representing lexical forms to concrete objects in the world. Studies of early spoken language comprehension have measured children’s gaze as they look at pairs of familiar objects while listening to speech naming one of the objects (Bergelson & Swingley, 2013; Fernald, Pinto, Swingley, Weinberg, & McRoberts, 1998). Such research shows that young listeners process language incrementally, shifting their gaze to a named object as soon as the auditory information is sufficient to enable referent identification. Moreover, individual differences in real-time processing efficiency predict vocabulary growth and later language and cognitive outcomes (Fernald, Perfors & Marchman, 2006; Marchman & Fernald, 2008).

However, little is known about how young children learning a visual language develop skill in processing signs from moment to moment. Here we ask whether children learning American Sign Language (ASL) develop skill in real-time processing of signs in ways that are parallel to children learning spoken language, and whether variability among ASL-learning children in real-time processing skills are related to their expressive vocabulary development, as in children learning spoken language. We also explore how deaf children learning ASL compare with hearing children learning ASL in the accuracy and time course of real-time lexical processing.

### ASL processing in adults

Psycholinguistic studies of adults show that language processing in signed and spoken languages is similar in many ways. For example, as in spoken language processing, signers are influenced by both lexicality and frequency; non-signs are identified more slowly than real signs (Corina & Emmorey, 1993) and high frequency signs are recognized faster than low frequency signs (Carreiras, Gutiérrez-Sigut, Baquero, & Corina, 2008). Using an eye-tracking procedure, Lieberman, Borovsky, Hatrak, & Mayberry (2014) found that adult signers are also sensitive to sub-lexical features of signs during real-time comprehension, showing evidence of incremental semantic processing.

However, differences between spoken and signed languages in the linguistic structure and surface features of lexical forms could have consequences for the time course of sign interpretation (Corina & Knapp, 2006). Using a gating procedure, Emmorey & Corina (1990) showed deaf adults increasingly longer videos of signs in isolation and asked them to identify the signs in an open-ended, non-timed response format, while English speakers heard increasingly longer segments of spoken words in isolation. Accurate identification of signs required relatively less of the linguistic signal as compared to spoken word identification, suggesting that features of visual-manual languages such as simultaneous presentation of phonological information might alter the time course of lexical access. Thus, there is evidence of both parallels and differences between signed and spoken language processing by adults. However, no previous studies have explored the development of real-time language comprehension in young ASL-learning children.

### Lexical development in ASL

Diary studies of sign language acquisition show that ASL-learning children follow a similar developmental path as children learning spoken languages (Lillo-Martin, 1999; Mayberry & Squires, 2006). For example, young signers typically produce recognizable signs before the end of the first year and two-sign sentences by their 2nd birthday (Newport & Meier, 1985). And as in spoken language, young ASL learners tend first to learn more nouns than verbs or other predicates (Anderson & Reilly, 2002).

Other research has investigated how the visual nature of sign language might influence children’s interactions with caregivers and thus also affect learning mechanisms such as joint attention that support lexical development (e.g., Tomasello & Farrar, 1986). Because children learning ASL must rely on vision both to process linguistic information and to look at referenced objects, they must alternate gaze between the signer and objects in the environment to achieve joint attention (Waxman & Spencer, 1997(?); Harris & Mohay, 1997). In a study comparing gaze patterns in deaf and hearing caregiver-child dyads, Lieberman, Hatrak, & Mayberry (2014) found that deaf children frequently shifted their gaze to the caregiver during book reading in order to maintain contact with the signed signal. Hearing children, in contrast, looked continuously at the book while the caregiver was speaking, rarely shifting gaze to the caregiver. However, no prior work has compared the gaze behavior of deaf and hearing children learning ASL. One possibility is that the time course of lexical access will be similar for both groups, driven by their experience with a visual language. Another possibility is that deaf children will be slower to disengage from a signer, since they must rely on vision to monitor additional linguistic signal. In the current work, we directly test these possibilities by comparing the time course of real-time processing for deaf and hearing children learning ASL,

Taken together, these findings show that lexical development in children learning signed and spoken languages is parallel in important ways, but that modality-specific features make control of gaze potentially more important for children learning a visual language. By examining these critical visual attention skills in both deaf and hearing signers, this study is the first to reveal links between early individual differences in real-time language comprehension skills and other language outcomes in visual language learners.

### Research questions

By adapting a well-established paradigm for measuring spoken language processing efficiency for use with young visual language learners, this study addresses three main questions. First, do children learning ASL develop skill in real-time processing of familiar signs in ways that are parallel to children learning spoken language? Second, are differences among ASL-learning children in their real-time processing skills related to differences in expressive vocabulary development, as in children learning spoken language? And third, how do deaf children learning ASL compare with hearing children learning ASL in the accuracy and time course of real-time lexical processing?

# Method

### Participants

Participants were 16 deaf and 13 hearing children with native exposure to ASL (17 females, 12 males, = 28.5 months, range = 16-53 months) and 19 fluent adult signers, recruited by bi-cultural/bilingual researchers fluent in ASL. Children learning ASL from birth from a native signer are a difficult population to recruit, given that approximately 95% of deaf children are born to hearing parents with little prior exposure to a signed language (Mitchell & Karchmer, 2004). All participants were exposed to ASL from birth through extensive interaction with at least one fluent ASL caregiver, and they currently used ASL as their primary mode of communication at home. The majority of children also attended a center-based early childhood education program in which ASL was the primary mode of instruction. An additional 17 participants were tested but not included in the analyses because they were not exposed to ASL from birth (*n* = 12) or they did not complete the VLP task (*n* = 5).

### Measures

*Expressive vocabulary size*: Parents completed a 90-item vocabulary checklist based on the MacArthur-Bates Communicative Development Inventories (Fenson et al., 2007) and designed to be linguistically appropriate for children learning ASL. Vocabulary size was computed as the number of signs reported to be produced.

*ASL Processing*: Efficiency in online comprehension was assessed using a version of the looking-while-listening procedure (Fernald et al., 2006) adapted for ASL learners, which we call the Visual Language Processing (VLP) task[[1]](#footnote-1). Since this is the first study to measure online ASL processing efficiency in children of this age, several important modifications to the procedure were made, which we describe below.

### Stimuli

On each trial, the stimuli consisted of a pair of familiar objects with a central video of an adult female signing the name of one of the pictures. To allow for generalization beyond characteristics of a specific signer and sentence structure, ASL sentences were videorecorded from two native ASL users who used two different but acceptable ASL sentence structures for asking questions[[2]](#footnote-2):

* Sentence-initial wh-phrase: “HEY! WHERE [target noun]?”
* Sentence-final wh-phrase: “HEY! [target noun] WHERE?”

Before each sentence, the signer used a hand-wave gesture commonly used in ASL discourse to gain an interlocutor’s attention before initiating a linguistic utterance. This served to draw the children's attention to the signer in preparation for the upcoming sentence.

Target nouns consisted of eight object names familiar to most children learning ASL at this age. Visual stimuli consisted of colorful digitized pictures of these objects presented in four fixed pairs, in which the object names had no phonological overlap (cat—bird, car—book, bear—doll, ball—shoe). To prepare the stimuli, two female native ASL users recorded several tokens of each sentence in a child-directed register. These candidate stimuli were digitized, analyzed, and edited using Final Cut Pro software. The final tokens were chosen based on naturalness. Five filler trials were interspersed among the 32 test trials (e.g. “YOU LIKE PICTURES? MORE WANT?”). Images were digitized pictures presented in fixed pairs, matched for visual salience with 3–4 tokens of each object type. Each object was a target four times and a distractor four times. Finally, the side of target picture was counterbalanced across trials.

### Apparatus and Trial Structure

In the VLP task, stimuli were presented using a Macbook Pro laptop with a 27” monitor. The child sat on the caregiver’s lap, and the child’s gaze was recorded using a digital camcorder set up behind the monitor. To minimize visual distractions, testing occurred in a portable 5’ by 5’ tent with opaque walls, which reduced the potential for visual distractions during the task.

Figure 1 shows the structure of a trial with one question type (sentence final wh-phrase) in the VLP task. On each trial, the child saw two images of familiar objects on the screen for 2 s before the signer appeared. This allowed the child to inspect both images prior to the start of the sentence. Next, children saw a still frame of the signer for 1 s, which gave them the opportunity to orient to the signer prior to sentence onset. The target sentence was then presented, followed by a question and 2-s hold, followed by an exclamation to encourage attention to the task. Each trial lasted approximately 7 s.

**Figure 1:** Overview of the trial structure for one question type (sentence final wh-phrase) on the VLP task.

### Coding and reliability

Children’s and adults’ gaze patterns were videotaped and later coded frame-by-frame (33-ms resolution) by highly-trained coders blind to target side. On each trial, coders indicated whether the eyes were fixated on the central signer, one of the images (left or right picture), shifting between pictures, or away (off). This coding yielded a high-resolution record of eye movements aligned with target noun onset. Prior to coding, all sessions were pre-screened for parental interference and excluded on a trial-by-trial basis. To assess inter-coder reliability, 25% of the videos were re-coded. Agreement within a single frame averaged 98% on these reliability assessments.

### Calculating linguistic processing efficiency

*Computing target sign onset.* In the VLP task, computing accuracy and RT requires defining the appropriate response window, starting at the earliest moment when there is sufficient linguistic information to discriminate the pairs of pictures and to initiate a shift in gaze from the central signer to the named target image. In studies of spoken language processing, target word onset is typically identified using acoustic analysis software that measures the moment in the auditory signal when there first is acoustic evidence of the target noun. In signed languages like ASL, phonological information is presented simultaneously in several parts of the visual signal (e.g., hands and face) making it difficult to precisely determine the beginning of the target sign. In the VLP task, this problem is somewhat simplified because the pictures are presented in yoked pairs; thus on each trial, target sign onset is always determined in reference to a particular distracter object, the name of which does not overlap phonologically with the target picture. Thus, target sign onset can be defined as the earliest point in the signed sentence when the two pictures can be reliably discriminated.

Here, we took an empirical approach to defining target sign onset. As a starting point, the first and second authors, both fluent ASL signers, viewed each stimulus sentence and achieved a consensus regarding the onset of the target noun. The next step was to validate these preliminary judgments by asking fluent native signers to identify the target pictures based on videos in which different amounts of the target sign were shown. Thus, for each sign token, we created six videos, each showing a different amount of the target sign, ranging from three frames before to three frames after the noun onset determined based on our initial consensus judgments. Since our experimental stimuli consisted of 3-4 tokens for each of the 8 target nouns (28 tokens), fluent adult signers unfamiliar with the stimuli (*n* = 10) watched 168 videos (28 x 6 videos) while viewing the same picture pairs as in the VLP task. On each trial, the signers made forced-choice decisions indicating which of two images was signed in the video, yielding a proportion correct target identification for each of the six videos. For each sign token, final target noun onset values were identified as the earliest point in the signed sentence at which adults discriminated the pictures with 100% agreement.

*Reaction Time:* In the VLP task, four different types of responses are possible on a given trial: (1) signer-to-target object shift, (2) signer-to-distractor object shift, (3) signer-to-away shift, and (4) no shift. Reaction time (RT) corresponds to the latency to shift away from the central signer to the target picture on all signer-to-target shifts, measured from the empirically-defined onset of the target sign. Following Ratcliff (1993), we chose specific cutoff response times based on the empirical distribution of children’s RTs in our task. We selected the middle 90% of the RT distribution (600-2200 ms), since these data are the most likely to be generated by the underlying process of interest: children’s lexical access. Incorrect shifts (signer-to-distracter (19%), signer-to-away (14%), no shift (8%) were not included in the computation of median RT[[3]](#footnote-3). Since children varied in the likelihood that they would generate a signer-to-target shift, mean RTs were based on different numbers of trials across participants (*M* = 12.7 trials, range = 3—25).

*Target Accuracy:* Correct looking to the named target picture is a function of the child’s tendency to shift quickly from the central signer to the target picture in response to the target sign, and also to maintain fixation on the target picture. To determine the degree to which participants fixated the named picture across trials, mean proportion looking to target was calculated for each participant at each 33 ms frame from the onset of the target noun. Accuracy was defined as the mean proportion of time spent looking at the target picture out of the total time spent on either the target picture or the distracter picture from 600 to 2200 ms[[4]](#footnote-4) from target noun onset. Although all children and response types were included in the computation of accuracy, the number of trials contributing to the analysis varied across participants (*M* = 19.1).

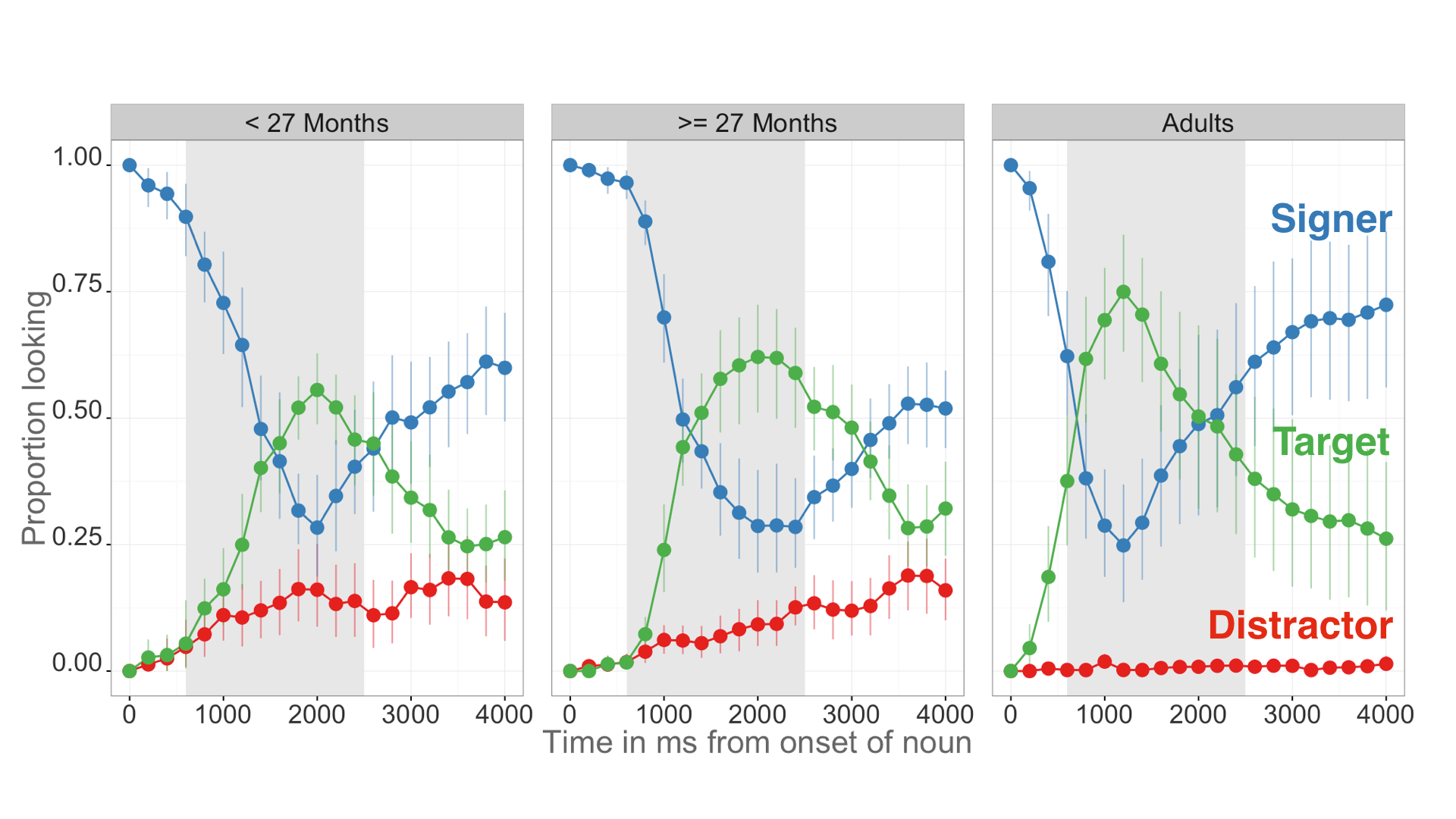
# Results

First, we present a high-level overview of our analytic approach and a qualitative analysis of developmental changes in looking behavior on the VLP task. Then, we present a series of Bayesian regression analyses showing that a) older children were faster and more accurate at comprehending familiar signs compared to younger children and b) children with better ASL processing skills also had larger ASL vocabularies. Finally, we compare the performance of deaf and hearing ASL-learning children and find no meaningful difference between the two groups, suggesting that auditory experience does not change the dynamics of reference in ASL.

We chose to use Bayesian analyses because it allowed us to include relevant prior knowledge about each participant in order to better estimate the strength of the associations between Reaction Time (RT) on the VLP task and age/vocabulary. Specifically, the use of RT as a processing measure is based on the assumption that the timing of children’s first shifts are generated by the speed of lexical access, and not the result of random guessing. Thus, we created an analysis model where participants who were more likely to be guessers would have less of an influence on the estimated relations between RT and age/vocabulary.

To quantify each participant’s probability of guessing, we computed the proportion of signer-to-target (correct) and signer-to-distracter (incorrect) shifts for each child. Previous work using the Looking-While-Listening paradigm could not easily compute these values, since the task did not include a center fixation point. We then used a latent mixture model in which we assumed that the observed data (children’s initial shifts away from the signer) were generated by two processes (guessing and knowledge) that had different overall probabilities of success, with the “guessing group” having a probability of 50% and the “knowledge” group having a probability > 50%. The group membership of each participant was a latent variable inferred based on that participant’s proportion of correct signer-to-target shifts relative to the overall proportion of correct shifts across all participants (see Lee & Wagenmakers [2013] for a detailed discussion of this modeling approach). We then used each participant’s inferred group membership to decide whether that participant was included in the linear regression predicting RT[[5]](#footnote-5). Importantly, this allowed us to include participants *proportional* to our belief that they were guessing, as opposed to simply removing them from the final analyses, thus maximizing the data that we had collected. It is important to point out that we use this approach only in the analysis of RT because we think that “guessing behavior” is part of the underlying process of interest when measuring children’s Accuracy on the VLP task.

In all of the Bayesian linear models[[6]](#footnote-6), we assume that each outcome variable (mean accuracies and reaction times for each participant) is drawn from a Gaussian distribution with a mean, μ, and a standard deviation, σ. The mean is generated by a linear function consisting of an intercept term, α, which encodes the expected value of the outcome variable when the predictor is zero, and a slope term, β, which encodes the expected change in the outcome with each unit change in the predictor (i.e., the strength of association). We use weak priors for the intercept and the standard deviation, allowing the model to consider a wide range of plausible values. For the prior on the slope parameter, we use a weak Gaussian distribution truncated at zero, encoding our directional hypotheses for the relations between processing skills and age/vocabulary (i.e., that we predict that these relations should be null or improve with increasing age and larger vocabulary size). For each analysis, we present a) the point estimates of α and β that maximize the posterior probability, and b) the 95% Highest Density Interval (HDI) of each parameter’s posterior distribution, which provides information about the uncertainty of the estimate[[7]](#footnote-7).

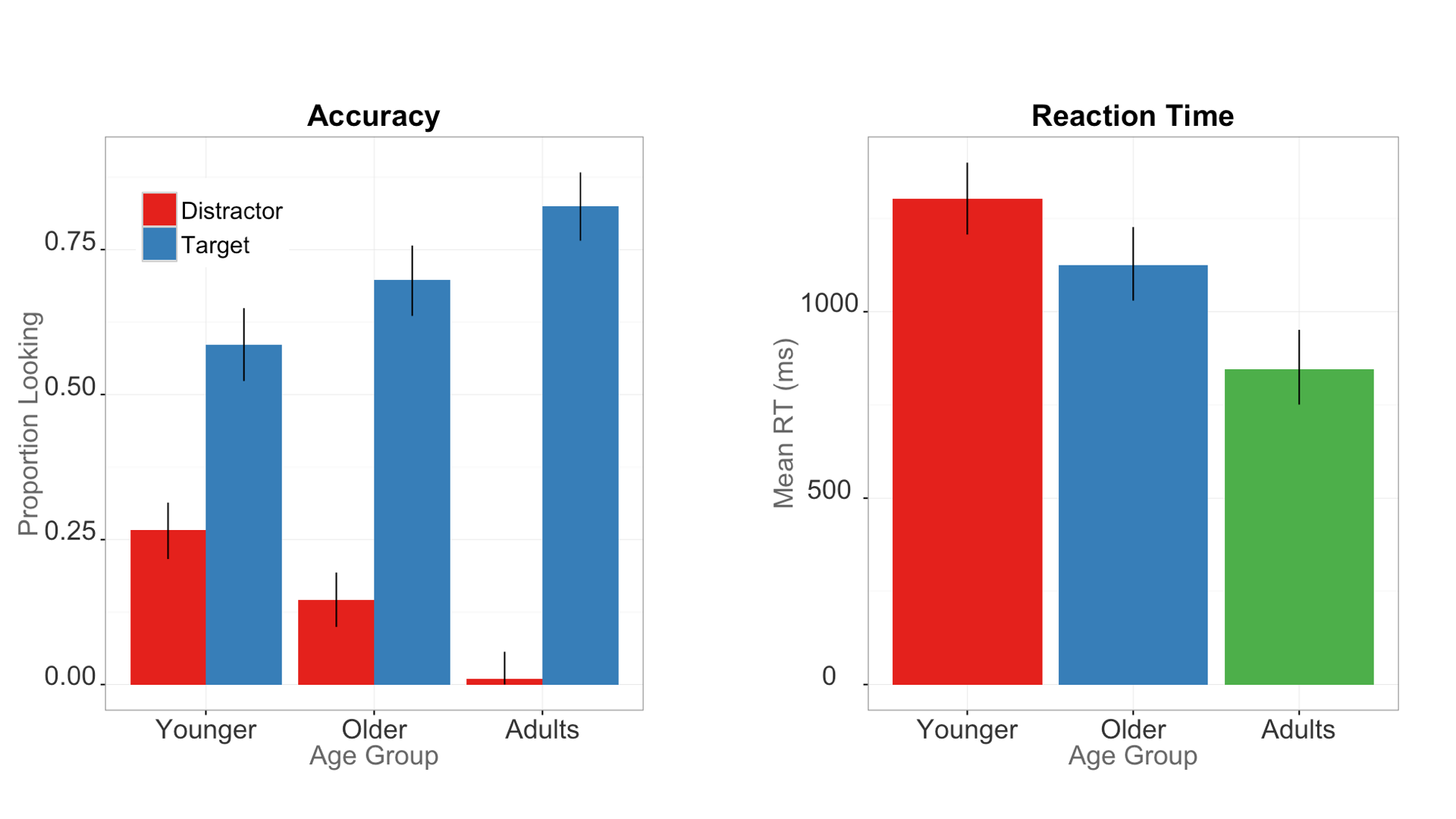
**Figure 2:** *An overview of the time course of looking behavior for younger children, older children, and adults. The curves show proportion looking to the signer (blue), the target image (green), and the distractor image (red). The grey shaded region represents the analysis window (600-2200ms) and the error bars represent +/- 95% CI computed by non-parametric bootstrap.*

### Overview of ASL processing

In the first set of analyses, we compare real-time processing skills by ASL-learning children and fluent adult signers. Figure 2 provides an overview of the time course of looking behavior in the VLP task[[8]](#footnote-8). The three curves show changes in the mean proportion of trials on which participants in each age group fixated the signer, the target image, or the distractor image at every 33 ms interval of the stimulus sentence. At the onset of the target sign, all participants were looking at the signer. As the sign unfolded, mean proportion looking to the signer decreased rapidly as participants shifted their gaze to the target or the distractor image. Proportion looking to the target increased sooner in the sentence and reached a higher asymptote compared to proportion looking to the distractor for all age groups. When modeling the difference between proportions looking to the target vs. the distractor pictures for each age group, all three groups spent more time looking at the target (younger:= 0.32, 95% HDI [0.24, 0.40]; older:**=** 0.55, 95% HDI [0.47, 0.62]; adults: **=** 0.81, 95% HDI [0.74, 0.89]). Moreover, when we model mean target looking as a function of age group, the 95% HDI for each group did not include the value of 0.5 (younger:**β=** 0.59, 95% HDI [0.52, 0.65]; older:**β=** 0.70, 95% HDI [0.64, 0.76]; adults: **β=** 0.82, 95% HDI [0.76, 0.88]), providing evidence that as a group even the youngest children were responding meaningfully to the linguistic signal. After looking to the target image, participants tended to rapidly shift their gaze back to the signer, reflected by the increase in proportion looking to signer around 2000 ms after target noun onset. In all three groups, proportion looking to distracter was small, decreasing to almost zero in the adults (Younger = 0.26, Older = 0.12, Adults = 0.05).

Figure 2 also provides a visual overview of age-related change in real-time ASL processing efficiency. Older children spent more time looking to the target picture compared to younger children, but not as much as adults (reflected by the increase in asymptote for the target looking curve across the three age groups). Moreover, older children tended to shift to the target picture sooner in the sentence than younger children, but not as rapidly as adults (reflected by the slope of the target looking curve).

Figure 3 shows group-level summary measures of ASL processing efficiency. Older children were more accurate than younger children ( = 0.11, 95% HDI [0.02, 0.19]) and had shorter mean latencies to orient to the target image ( = -181.98, 95% HDI [-40.01, -315.50]). As a group, children were less accurate ( = -0.18, 95% HDI [0.10, 0.26]) and slower to shift to the target image ( = 379.28, 95% HDI [254.18, 526.51]) compared to adults. Importantly, none of the HDIs for each parameter estimate included zero, providing evidence that we were able to reliably measure group-level changes in ASL processing efficiency.

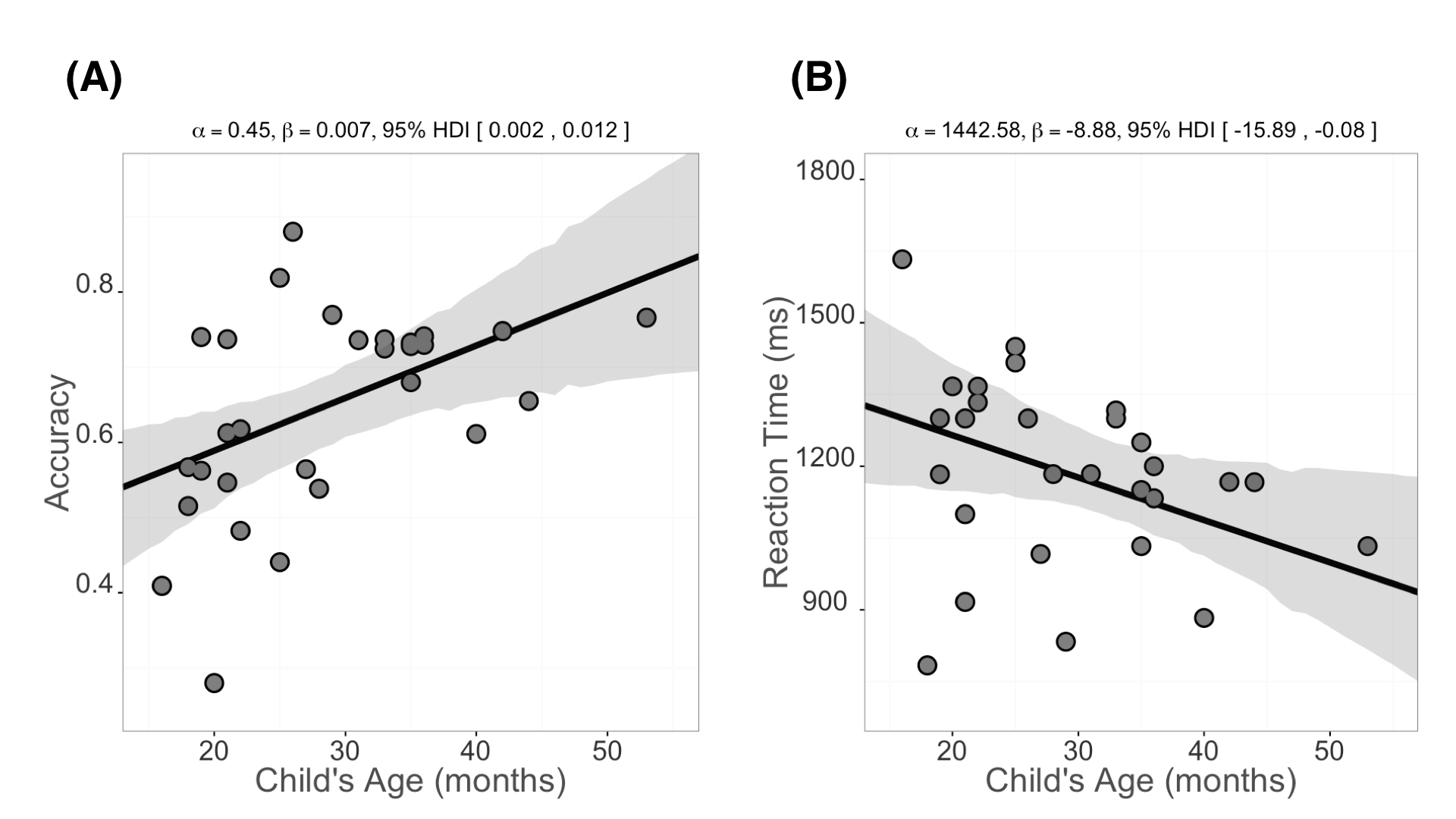
**Figure 3.** *Summary measures of developmental changes in ASL processing efficiency. The left panel shows mean Accuracy; the right panel shows mean RT. Error bars represent +/- 95% Highest Density Intervals.*

|  |  |  |  |
| --- | --- | --- | --- |
| Model | β MAP | 2.5% HDI | 97.5% HDI |
| **Accuracy ~ Age** | 0.007 | 0.002 | 0.012 |
| **Accuracy ~ Vocab** | 0.003 | 0.001 | 0.006 |
| **RT ~ Age** | -8.88 | -15.89 | -0.08 |
| **RT ~ Vocab** | -5.84 | -11.14 | -1.29 |

***Table 1:*** *Summary of the four univariate linear models using age and vocabulary size to predict accuracy and reaction time.* β *MAP is the maximum a posteriori estimate for the slope parameter in each model. The Highest Density Interval (HDI) shows the distribution of 95% of the plausible parameter values given the model and the data.*

### Links between processing efficiency and age

In the next set of analyses, we use Bayesian linear regressions to explore the quantitative relations between individual children’s age and real-time processing skills. Table 1 shows a summary of all four linear models that we report here. Mean accuracy scores, computed rover the 600–2200 ms window from noun onset, were positively associated with age (**β=** 0.007, 95% HDI [0.002, 0.012]), indicating that older ASL learners were more accurate than younger children in fixating the target picture. Quantitatively, our model estimates an increase of 0.7% points for each month of age, meaning that over the course of one year the model estimates a ~8% point gain in accuracy on the VLP task. Moreover, the value of zero was not included in the 95% Highest Density Interval, providing evidence for a positive association between age and accuracy.

***Figure 4:*** *Scatterplots of the relations between children’s age and measures of their accuracy (panel A) and reaction time (panel B) in the VLP procedure. Each point represents a participant. The solid black line is the maximum a posteriori model estimate for the mean accuracy at each age point. The shaded gray regions represent the 95% Highest Density Interval around the regression line.*

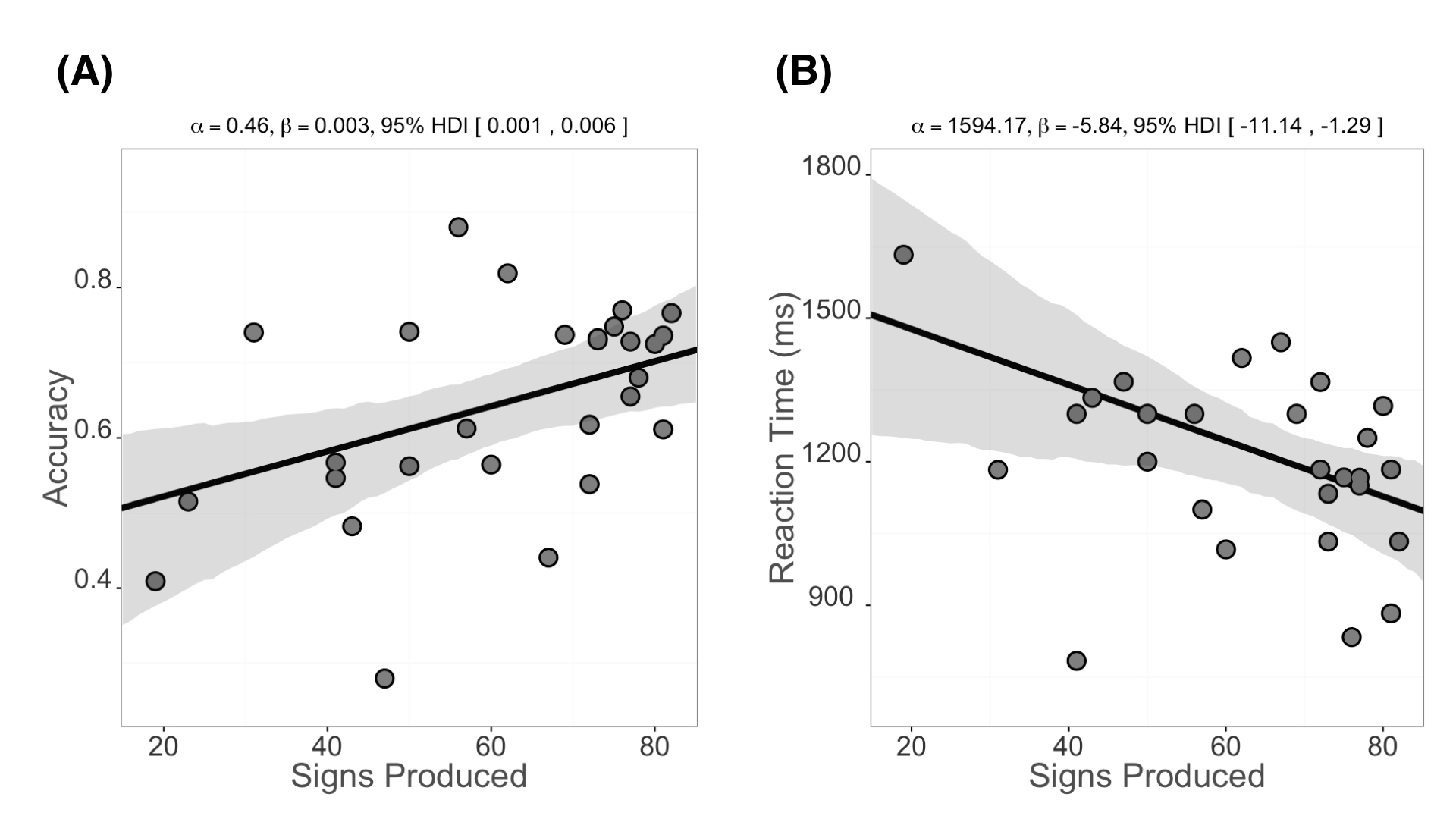
Mean reaction times were negatively associated with age (**β=** -8.88, 95% HDI [-15.89, -0.08]), with older children being faster to shift to the target picture compared to younger ones. The model estimates a ~9 ms gain in RT for each month, leading to ~108 ms gain over a year of development. However, the 95% HDI for the slope parameter in this analysis approaches zero, suggesting that there is relatively more uncertainty in the association between age and RT. Mean RTs were also related to mean accuracy scores such that those children who were faster to shift to the target were also more likely to maintain fixation on the target image throughout a greater proportion of the analysis window (**β=** -571.73, 95% HDI [-1357.60, -0.09]).

Together, the accuracy and reaction time analyses show that signers will reliably leave a central signer to shift to a target image in the VLP task. Importantly, signers varied in their response times and accuracy, and this variation was meaningfully linked to age. Thus, like children learning spoken language, ASL learners improve their real-time language processing skills over the second and third years of life as they make progress towards adult levels of language fluency.

### Links between processing efficiency and vocabulary

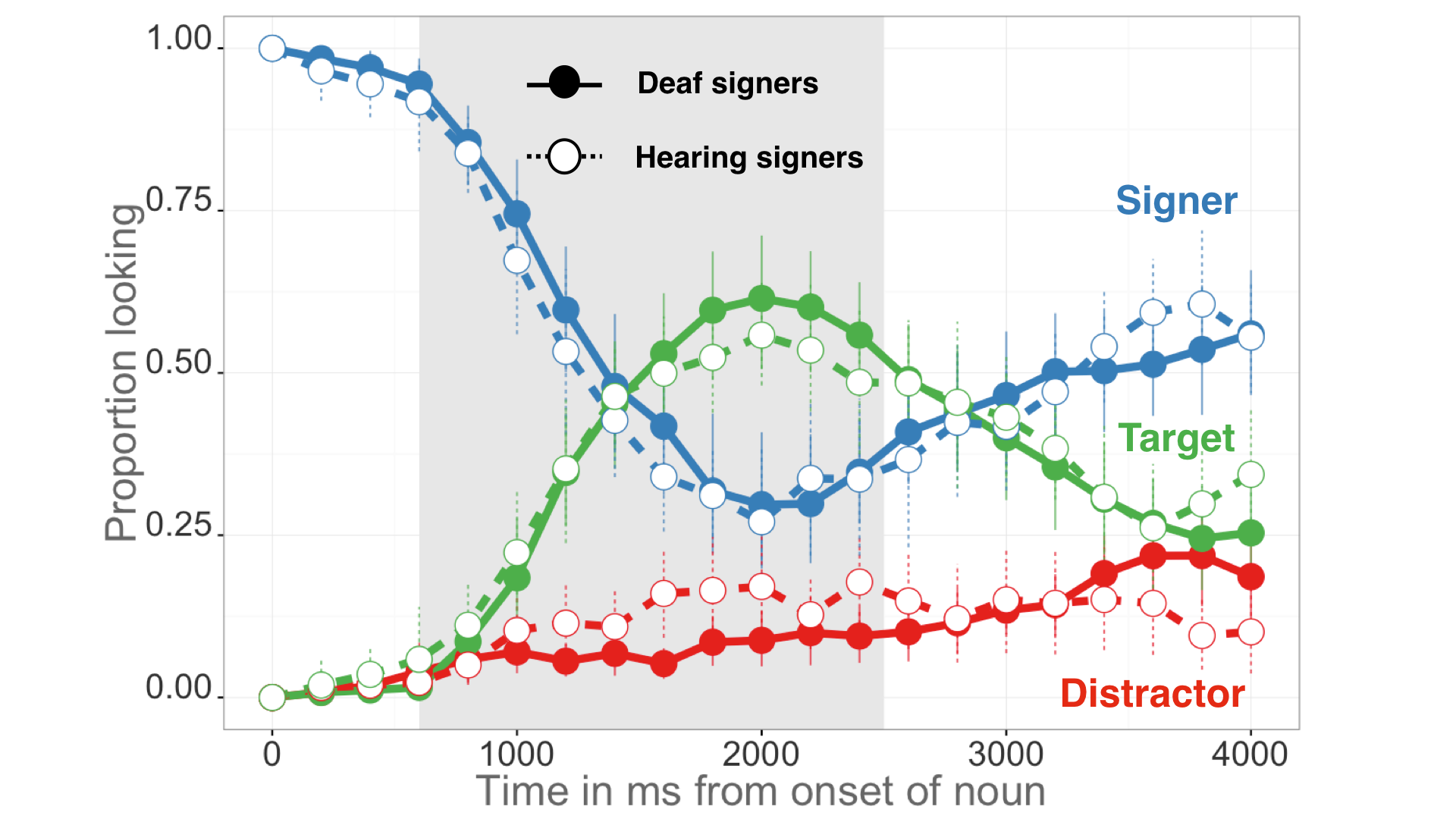
The next question we addressed was whether individual differences in processing skills were related to the size of children’s ASL vocabularies. Figure 5 shows relations between processing measures, accuracy and reaction time, and children's productive ASL vocabulary. Mean accuracy was positively related to vocabulary size (**β=** 0.003, 95% HDI [0.001, 0.006]) such that children with higher accuracy scores also had larger productive vocabularies, with the model estimating a 0.3% point gain for each additional sign children knew.

Moreover, mean reaction times were negatively correlated with vocabulary (**β=** -5.84, 95% HDI [-11.14, -1.29]), indicating that children who were faster to recognize ASL signs were those with larger sign vocabularies, with each additional sign resulting in a ~6 ms decrease in estimated RT. Importantly, while there was relatively more uncertainty in the association between age and RT, we did not see the same uncertainty in the estimate of the link between vocabulary and RT.



**Figure 5:**  Scatterplot of relations between children’s productive ASL vocabulary and measures of their accuracy (panel A) and reaction time (panel B) in the VLP procedure. Plotting conventions are the same as in Figure 4.

It is important to point out that age and vocabulary were strongly correlated in our sample ( = 0.74). Thus, when you include both predictors in a multiple regression analysis predicting either RT or Accuracy, the parameter values for each predictor variable tradeoff such that there are a wide range of plausible values that could explain the data. This makes it difficult to make inferences about the unique contribution of each predictor variable (see McElreath [2016] for a discussion of issues related to multicollinearity in this modeling approach). However, this study was not designed to measure the separate contribution of vocabulary or age. But based on previous work (e.g., Fernald & Marchman, 2012), we think that if we were to collect data from a narrow age range (experimentally controlling for age), then vocabulary would emerge as the strongest predictor of ASL processing skills.

******Taken together, these analyses indicate that older children with larger expressive vocabularies were more accurate and efficient in identifying the referents of familiar signs.. These findings parallel the now large body of previous research with monolingual children learning English or Spanish (Fernald et al., 2006; Hurtado, Marchman, & Fernald, 2007).

***Figure 6.*** *The time course of looking behavior for deaf and hearing signers. The curves show proportion looking to the signer (blue), the target image (green), and the distractor image (red). The circle fill and the line type represent hearing status; the grey shaded region represents the analysis window (600-2200ms); error bars represent +/- 95% CI computed by non-parametric bootstrap.*

### Effects of hearing status

In the final set of analyses, we compared deaf and hearing native ASL learning children’s performance on the VLP task. This exploratory analysis allows us to ask if auditory experience changes the dynamics of reference in a visual language. Figure 6 shows an overview of looking behavior in the VLP task for deaf (n=16, = 28m, = 7.48m) and hearing (n=13, = 29m, = 11.19m) children. Overall, these two groups showed a similar time course of looking behavior: shifting away from the signer, increasing looks to the target, and shifting back to the signer at similar time points as the sign unfolded. We found no differences in Accuracy (= -0.04, 95% HDI [-0.12, 0.04]) or Reaction Time (= 69.25, 95% HDI [-83.60, 237.21]), with the HDI including zero for both models. These analyses provide evidence that auditory experience does not qualitatively change patterns of visual language processing observed in the VLP task, suggesting that experience with a visual language drives children’s performance.

# Discussion

Establishing reference in real-time is a fundamental component of language learning. With this study, we aimed to develop and validate the first measures of young ASL learners’ real-time language comprehension skills and to explore how language processing skills are linked to age, vocabulary, and hearing status. There are three main findings from this research.

The first main finding is that, like children learning spoken language (Fernald et al., 1998), young ASL learners' showed measurable age-related improvement in the efficiency with which they processed language. Even in ASL-learning 2-year olds, shifting from the signer to the target picture occurred rapidly, with few false alarms to the distracter. All of the target signs were familiar to children; yet when compared to the younger children, older children identified the correct referents more quickly and accurately and were less likely to fixate the unlabeled picture. These developmental changes provide additional evidence that children learning ASL in native contexts follow a similar developmental path to that of children learning spoken language (Lillo-Martin, 1999; Mayberry & Squires, 2006). Moreover, this is the first study to show that real-time ASL *comprehension* skills are linked to early linguistic development. Prior work on the developmental trajectories of deaf children has relied on aspects of language production, often because production is easy to observe, and thus is easier to measure than comprehension. However, it is a well-known phenomenon in language acquisition that comprehension tends to precede production (see Clark, 2009). By developing a precise measure of real-time ASL comprehension, we have provided a tool that enables us to study the emergence of children's language skills earlier in development.

The second main result is the discovery of a link between early ASL processing skills and children's productive ASL vocabularies. ASL-learning children who knew more signs were also faster and more accurate in identifying the correct referent than those who were lexically less advanced. These results with children learning ASL are consistent with other studies with English- and Spanish- learning children, which find strong relations between efficiency in online language comprehension and other concurrent and longitudinal measures of linguistic achievement (Fernald et al., 2006, 2001; Marchman & Fernald, 2008; Zangl et al., 2005).

The third finding is that deaf and hearing children learning ASL as a first language showed similar patterns of visual language processing. Both groups showed similar speed of processing and spent about the same amount of time looking to the target image before looking back to the signer. These parallels suggest that even though hearing children can use both vision and hearing to process incoming information, this experience does not qualitatively change the time course of visual language processing, which instead appears to be linked to exposure to a visual language.

### Limitations

This research has several limitations. First, while the sample size in this study was large relative to those in previous research on early ASL development, it was still a small sample, and the results should be replicated in future studies that include more participants. We have made all of our stimuli, data, and analysis code publicly available at the project page for this study[[9]](#footnote-9), with the hope that other researchers will benefit from what we have learned in this work.

Second, our sample included young children across a broad age range. Recall that using both age and vocabulary to predict accuracy or reaction time resulted in high levels of uncertainty about the contribution of either predictor. Based on the strength of past evidence, it is likely that testing groups of children in narrower age ranges would allow us to see independent effects of vocabulary size on both ASL processing measures. Thus more evidence is needed in order to best characterize the relations between accuracy, RT, and vocabulary in children learning ASL.

Third, the novelty of the VLP task makes it difficult to directly compare our findings with previous work on ASL and spoken language processing. For example, in contrast to prior ASL gating studies with adults (e.g., Emmorey & Corina, 1990; Morford & Carlsen, 2011), our stimuli were full sentences signed in a child-directed register, not isolated signs, and our dependent measure was the latency to shift gaze away from a signer to one of two images, not a free response. In addition, the VLP task includes a central fixation image – the signer – making it substantially different in task demands from previous studies of the development of children’s spoken language processing (e.g., Fernald, 1998). These differences do not allow us to make any general claims about the time course of processing in signed vs. spoken languages in absolute terms. Nevertheless, our results show impressive similarities with previous findings with learners of spoken languages in terms of age-related changes and links to measures of vocabulary.

Finally, our sample is not representative of the majority of children learning ASL in the United States. We took great care to include only children who are native signers with exposure to ASL from birth. However, we might anticipate that the development of real-time language processing might look different in children who are late learners or who have more heterogeneous and inconsistent language exposure. An important next step is to explore how differences among children in ASL processing skills are influenced by differences in their experience with signed languages. It is well known that children's efficiency of real-time processing of spoken language is linked to the quantity and quality of the speech that they hear (Hurtado et al., 2008; Weisleder & Fernald, 2013). We would expect similar relations in children learning ASL.

In sum, this study provides the first evidence that young ASL learners’ processing skills undergo developmental changes that are related to meaningful vocabulary outcomes. Such links between early processing efficiency and vocabulary contribute to the now significant body of literature highlighting parallels between signed and spoken language development when children are exposed to native sign input. Moreover, we found similar results for deaf and hearing ASL-learning children, suggesting that exposure to ASL, and not the experience of relying on vision to process both language and the visual world, shapes these real-time processing skills. We hope that the VLP task will provide a useful method for researchers and educators, providing a way to track developmental trajectories of children learning ASL.

# References (count = 33; limit = 40)

Anderson, D., & Reilly, J. (2002). The MacArthur communicative development inventory: Normative data for American Sign Language. *Journal of Deaf Studies and Deaf Education*, 83–106.

Arendsen, J., Van Doorn, A. J., & de Ridder, H. (2009). When do people start to recognize signs?. *Gesture*, *9*(2), 207-236.

Bergelson, E., & Swingley, D. (2012). At 6 to 9 months, human infants know the meanings of many common nouns. *Proceedings of the National Academy of Sciences of the USA, 109,* 3253-3258.

Carreiras, M., Gutiérrez-Sigut, E., Baquero, S., & Corina, D. (2008). Lexical processing in Spanish Sign Language (lSE). *Journal of Memory and Language*, *58*(1), 100–122.

Clark, E. V. (2009). *First language acquisition*. Cambridge University Press.

Corina, D. P., & Emmorey, K. (1993). Lexical priming in American Sign Language. In *34th annual meeting of the Psychonomics Society*.

Corina, D. P., & Knapp, H. P. (2006). Lexical retrieval inAmerican Sign Language production. *Papers in Laboratory Phonology*, *8*, 213–240.

Emmorey, K., & Corina, D. (1990). Lexical recognition in sign language: Effects of phonetic structure and morphology. *Perceptual and Motor Skills*, *71*(3f), 1227–1252.

Fenson, L. (2007). *MacArthur-Bates communicative development inventories: User’s guide and technical manual*. Paul H. Brookes Publishing Company.

Fernald, A., & Marchman, V. A. (2012). Individual differences in lexical processing at 18 months predict vocabulary growth in typically developing and late-talking toddlers. *Child Development*, *83*(1), 203–222.

Fernald, A., Perfors, A., & Marchman, V. A. (2006). Picking up speed in understanding: Speech processing efficiency and vocabulary growth across the 2nd year. *Developmental Psychology*, *42*(1), 98.

Fernald, A., Pinto, J. P., Swingley, D., Weinberg, A., & McRoberts, G. W. (1998). Rapid gains in speed of verbal processing by infants in the 2nd year. *Psychological Science*, *9*(3), 228–231.

Grosjean, F. (1980). Spoken word recognition processes and the gating paradigm. *Perception & Psychophysics*, *28*(4), 267–283.

Harris, M., & Mohay, H. (1997). Learning to look in the right place: A comparison of attentional behavior in deaf children with deaf and hearing mothers. *Journal of Deaf Studies and Deaf Education*, *2*(2), 95–103.

Hurtado, N., Marchman, V. A., & Fernald, A. (2007). Spoken word recognition by Latino children learning Spanish as their first language. *Journal of Child Language*, *34*(02), 227–249.

Hurtado, N., Marchman, V. A., & Fernald, A. (2008). Does input influence uptake? Links between maternal talk, processing speed and vocabulary size in Spanish-learning children. *Developmental Science*, *11*(6), F31–F39.

Lee, M., & Wagenmakers, E. (2013). Bayesian modeling for cognitive science: A practical course. *Cambridge UP*.

Lieberman, A. M., Borovsky, A., Hatrak, M., & Mayberry, R. I. (2014). Real-time processing of ASL signs: Delayed first language acquisition affects organization of the mental lexicon. *Journal of Experimental Psychology: Learning, Memory, and Cognition*.

Lieberman, A. M., Hatrak, M., & Mayberry, R. I. (2014). Learning to look for language: Development of joint attention in young deaf children. *Language Learning and Development*, *10*(1), 19–35.

Lillo-Martin, D. (1999). Modality effects and modularity in language acquisition: The acquisition of American Sign Language. *Handbook of Child Language Acquisition*, *531*, 567.

Marchman, V. A., & Fernald, A. (2008). Speed of word recognition and vocabulary knowledge in infancy predict cognitive and language outcomes in later childhood. *Developmental Science*, *11*(3), F9–F16.

Marslen-Wilson, W., & Zwitserlood, P. (1989). Accessing spoken words: The importance of word onsets. *Journal of Experimental Psychology: Human Perception and Performance*, *15*(3), 576.

Mayberry, R. I., & Squires, B. (2006). Sign language acquisition.

McElreath, R. (2016). Statistical Rethinking: A Bayesian Course with Examples in R and Stan (Vol. 122). CRC Press.

Meier, R. P., Mauk, C., Mirus, G. R., & Conlin, K. E. (1998). Motoric constraints on early sign acquisition. *29*, 63–72.

Mitchell, R. E., & Karchmer, M. A. (2004). Chasing the mythical ten percent: Parental hearing status of deaf and hard of hearing students in the United States. *Sign Language Studies*, *4*(2), 138–163.

Morford, J. P., & Carlson, M. L. (2011). Sign perception and recognition in non-native signers of ASL. *Language learning and development*, *7*(2), 149-168.

Newport, E. L., & Meier, R. P. (1985). *The acquisition of American Sign Language.* Lawrence Erlbaum Associates, Inc.

Petronio, K., & Lillo-Martin, D., (1997). WH-Movement and the Position of Spec-CP: Evidence from American Sign Language. *Language*, *73*(1), 18–57.

Plummer, M. (2003, March). JAGS: A program for analysis of Bayesian graphical models using Gibbs sampling. In *Proceedings of the 3rd international workshop on distributed statistical computing* (Vol. 124, p. 125). Technische Universit at Wien.

Ratcliff, R. (1993). Methods for dealing with reaction time outliers. *Psychological bulletin*, *114*(3), 510.45, R. P., & Spencer, P. E. (1997). What mothers do to support infant visual attention: Sensitivities to age and hearing status. *Journal of Deaf Studies and Deaf Education*, *2*(2), 104–114.

Weisleder, A., & Fernald, A. (2013). Talking to children matters: Early language experience strengthens processing and builds vocabulary. *Psychological Science*, *24*(11), 2143–2152.

Zangl, R., Klarman, L., Thal, D., Fernald, A., & Bates, E. (2005). Dynamics of word comprehension in infancy: Developments in timing, accuracy, and resistance to acoustic degradation. *Journal of Cognition and Development*, *6*(2), 179–208.

1. The stimuli can be viewed at the project page for this experiment: <https://github.com/kemacdonald/SOL>. [↑](#footnote-ref-1)
2. See Petronio, K. and Lillo-Martin, D., (1997) for a detailed discussion of the acceptability of these two question structures. [↑](#footnote-ref-2)
3. We chose to use the median because this point estimate is less sensitive to outliers, which can have a large effect on individual RT estimates when participants contribute a small number of RTs. [↑](#footnote-ref-3)
4. We use the same analysis window as in the RT analyses (the middle 90% of the RT distribution), since these data are the most likely to be generated by children’s lexical access. [↑](#footnote-ref-4)
5. Four children (ages: 18, 20, 22, and 25 months) had high posterior probability mass on guessing: posterior probabilities of 0.99, 0.98, 0.98, and 0.94 respectively (mean proportion signer-to-target scores for these participants were: 0.54, 0.58, 0.42, 0.29, 0.33). [↑](#footnote-ref-5)
6. All models were implemented in JAGS (Plummer, 2003). See the supplementary materials for details about model specifications and simulations. [↑](#footnote-ref-6)
7. The HDI can be interpreted as meaning there is a 95% chance that the true parameter value falls within this interval given the model specification and the data. [↑](#footnote-ref-7)
8. Preliminary analyses examined response patterns for the two sentence types separately and found no significant differences. Thus, all reported analyses collapse across the two sentence structures. [↑](#footnote-ref-8)
9. <https://github.com/kemacdonald/SOL> [↑](#footnote-ref-9)