Online language comprehension in children learning American Sign Language (ASL)

Kyle MacDonald

Todd LaMarr

David Corina

Virginia A. Marchman

Anne Fernald

**Word count: 2847** (acknowledgements, introduction, and discussion)

# Author note:

We are grateful to the many children and parents who participated in this research, and to the staff of the California School for the Deaf in Fremont. We extend a special thanks to Karina Pedersen and Lisalee Egbert for their help with participant recruitment, to Pearlene Utley and Rosa Lee Timm for help with creation of the stimuli, and to Shane Blau, Kat Adams, Melanie Ashland, and the staff of the Language Learning Lab at Stanford University. This work was made possible by an NIDCD grant to Anne Fernald and David Corina (R21 DC012505-1).

**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, for children learning a visual language like American Sign Language (ASL), eye movements and gaze are used to process *both* the linguistic signal and relevant information in the visual world to which the linguistic signal is referring. Can gaze patterns by ASL learners also provide an index of their lexical processing skills? In this study, we developed the first measures of children’s real-time ASL comprehension abilities, investigating links between these skills, children’s age, vocabulary, and hearing status. Native ASL-learning children (16-53 mos, *n*=29, 16 deaf and 13 hearing) and fluent adult signers (*n*=19) participated 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 significantly correlated 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 visual language processing skills are driven by experience with a visual language, and not by deafness. These novel findings show striking parallels between the development of language comprehension in visual language learners and in children learning spoken languages.

# Introduction

Understanding language rapidly and accurately is central to our ability to function effectively in daily life. A fundamental component of language understanding is establishing reference during *real-time* language interaction. Research with children learning spoken language shows that the ability to link abstract symbols (i.e., words) to concrete objects in the world undergoes significant developmental changes during the 2nd and 3rd years of life (Fernald, Perfors & Marchman, 2006). This work has used measures of children’s gaze shifts during sentence processing to provide a precise index of efficiency of spoken language processing, showing that listeners process spoken language *incrementally*, shifting their gaze off of the speaker as soon as there is sufficient auditory information to identify a referent. Moreover, individual differences in processing efficiency predict growth in vocabulary and later language and cognitive outcomes (Fernald & Marchman, 2012; Marchman & Fernald, 2008). In this study, we explore whether gaze shifts can reveal similar relations to development and vocabulary outcomes in children learning American Sign Language (ASL), a language presented in the visual modality.

Since the foundational work of Bellugi (1979) established that signed languages are natural human languages not derivative from spoken languages, researchers have explored similarities and differences in the acquisition of signed, compared to spoken languages. There are at least two reasons why gaze shifts might reflect different processes during spoken vs. visual language comprehension. Since children learning ASL must rely on vision both to process the incoming linguistic information and to look at objects in the visual scene, moving one’s eyes to identify the referent of a sentence may come at a higher cost for children learning ASL. To visually fixate a named object, children must disengage their eyes from the source of linguistic information, losing access to key portions of the signal. This tension creates a scenario where subsequent linguistic information might be missed. In addition, gaze serves other complex linguistic functions in ASL, such as regulating turn taking during conversation (Baker & Padden, 1978) and monitoring role shifts during narrative production (Bahan & Supalla, 1995). Thus, the multi-purpose and complex nature of gaze in ASL potentially undermines the assumption that eye-movements directly reflect children’s ASL comprehension skills.

In the current work, we investigate this assumption and ask whether gaze shifts are meaningfully related to efficiency of language comprehension in children learning ASL. Using a real-time language-processing task adapted for visual language learners, we explore age-related changes in ASL processing efficiency and test whether individual differences link to vocabulary development in deaf and hearing children learning ASL.

### Spoken language processing

To follow a typical conversation, skilled listeners must rapidly apprehend meaning in combinations of words from moment to moment as the speech signal unfolds at rates of 10-15 phonemes/second. Extensive research with adults using online measures shows that skilled listeners can identify spoken words with only partial phonetic information, evaluating hypotheses about word identity incrementally based on what they have heard up to that moment (Marslen-Wilson & Zwitserlood, 1989). Moreover, adults are adept in the parallel processing of multiple *cross-modal* streams of information, rapidly integrating the acoustic speech signal as it unfolds in time with information from the visual scene to derive intended meaning (Altmann & Kamide, 1999; Dahan & Tanenhaus, 2004; Tanenhaus, Spivey-Knowlton, Eberhard, & Sedivy, 1995).

Over the past fifteen years, research with infants and young children has incorporated the same high-resolution measures of language processing (Fernald, Pinto, Swingley, Weinberg, & McRoberts, 1998; Snedeker & Trueswell, 2004), making it possible to obtain continuous measures of speed and accuracy that enable sensitive assessment of efficiency in spoken language processing even by very young children. Using these procedures, researchers have found systematic age-related changes in the speed and accuracy of responses to familiar words (Fernald et al., 1998), and have shown that efficiency in word recognition is correlated with individual differences in vocabulary knowledge, as well as faster rates of vocabulary growth across the second year (Fernald, Perfors, & Marchman, 2006; Fernald & Marchman, 2012). The current study is the first to adapt these online processing efficiency measures for young children learning a signed language, ASL.

### ASL processing in adults

ASL is a visual-gestural language expressed with hands, arms and face, a modality difference with potential consequences for how linguistic information is processed. In many ways, adult language processing appears to be parallel in spoken and signed languages. Like spoken language users, signers show effects of: (a) lexicality, response times to identify non-signs are slower than for actual signs (Corina & Emmorey, 1993), (b) frequency, high frequency signs are recognized faster than low frequency signs (Carreiras, Gutiérrez-Sigut, Baquero, & Corina, 2008), and (c) phonological parameters, the sublexical units of sign – handshape, location, and movement – influence sign recognition (Carreiras et al., 2008; Corina & Emmorey, 1993; Hildebrandt & Corina, 2002). More recently, Lieberman, Borovsky, Hatrak, & Mayberry (2014a) found that early, but not late-learners, show activation of sub-lexical features during real-time comprehension and evidence of incremental semantic processing.

However, differences in linguistic structure and surface features of lexical forms in spoken vs. signed languages have consequences for the time-course with which signs are understood (Carreiras, 2010; Corina & Knapp, 2006). Using a gating procedure, Emmorey & Corina (1990) showed deaf participants increasingly longer videos of signs in isolation and asked them to identify the signs in an open-ended, non-timed response format. Analogously, English speakers heard different lengths of the acoustic signal. Sign identification was faster than word identification, suggesting that features of a visual-manual language (e.g., simultaneous presentation of phonological information) can 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 studies have explored the development of real-time language comprehension in young ASL-learning children.

### Lexical development in ASL

Diary studies of the acquisition of ASL in native-language contexts show that children follow a strikingly 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). Moreover, as in spoken language, young ASL learners tend to first acquire a preponderance of nouns, rather than verbs or other predicates (Anderson & Reilly, 2002).

A separate body of research has investigated how the visual nature of the language may have consequences for the mechanisms that support lexical development. For example, children learning ASL must alternate gaze between linguistic information and objects in real-world learning contexts to achieve joint attention (Waxman & Spencer, 1997). Harris & Mohay (1997) found that, at 18 months, deaf children frequently shifted visual attention towards their mothers during a free play interaction. More recent work by Lieberman, Hatrak, & Mayberry (2014b) showed that deaf children make frequent shifts in gaze during book reading in order to perceive both linguistic input and the non-linguistic context.

Taken together, these findings show that ASL development follows a similar trajectory as spoken language, and that precise control of gaze shifts are likely to be a critical component of ASL lexical development. However, no prior studies have systematically investigated how young deaf and hearing ASL learners establish reference during real-time sentence processing in a controlled experimental context. Such a study would provide insights into the development of these critical visual attention skills in young signers, filling an important gap in our understanding of whether early individual differences in language comprehension skills are linked to other meaningful linguistic outcomes in visual language learners.

### Current study

This study is the first to explore the early development of real-time processing of signs by very young children learning ASL. First, we adapt a well-established paradigm for measuring spoken language processing efficiency for use with young visual language learners. Next, we ask whether increased efficiency in early comprehension by native ASL learners follows a similar developmental trajectory as that of children learning spoken language. We then test whether individual variation in ASL processing skills show similar concurrent relations to children’s age and vocabulary size. Finally, we explore whether ASL processing is qualitatively similar for deaf and hearing children, asking whether auditory experience changes the time-course of children’s lexical access in ASL.

# Method

### Participants

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 were recruited from several locations by bi-cultural/bilingual researchers fluent in ASL. Native ASL learners are a difficult population to recruit because approximately 95% of deaf children are born to hearing parents with little prior exposure to a signed language (Mitchell & Karchmer, 2004). All children were exposed to ASL from birth via at least one fluent ASL caregiver and 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. Thus, all children were immersed in ASL, both at home and in the daycare setting. An additional 20 participants were tested, but not included in the analyses due to fussiness (n = 5), being outside the target age range (n = 3), or not receiving enough ASL exposure (n = 12). For visualization purposes, children were divided into two groups using a median split by age: Younger (< 27 Months), Older (>= 27 Months), but we conduct all critical statistical tests treating age as a continuous variable.[[1]](#footnote-1)

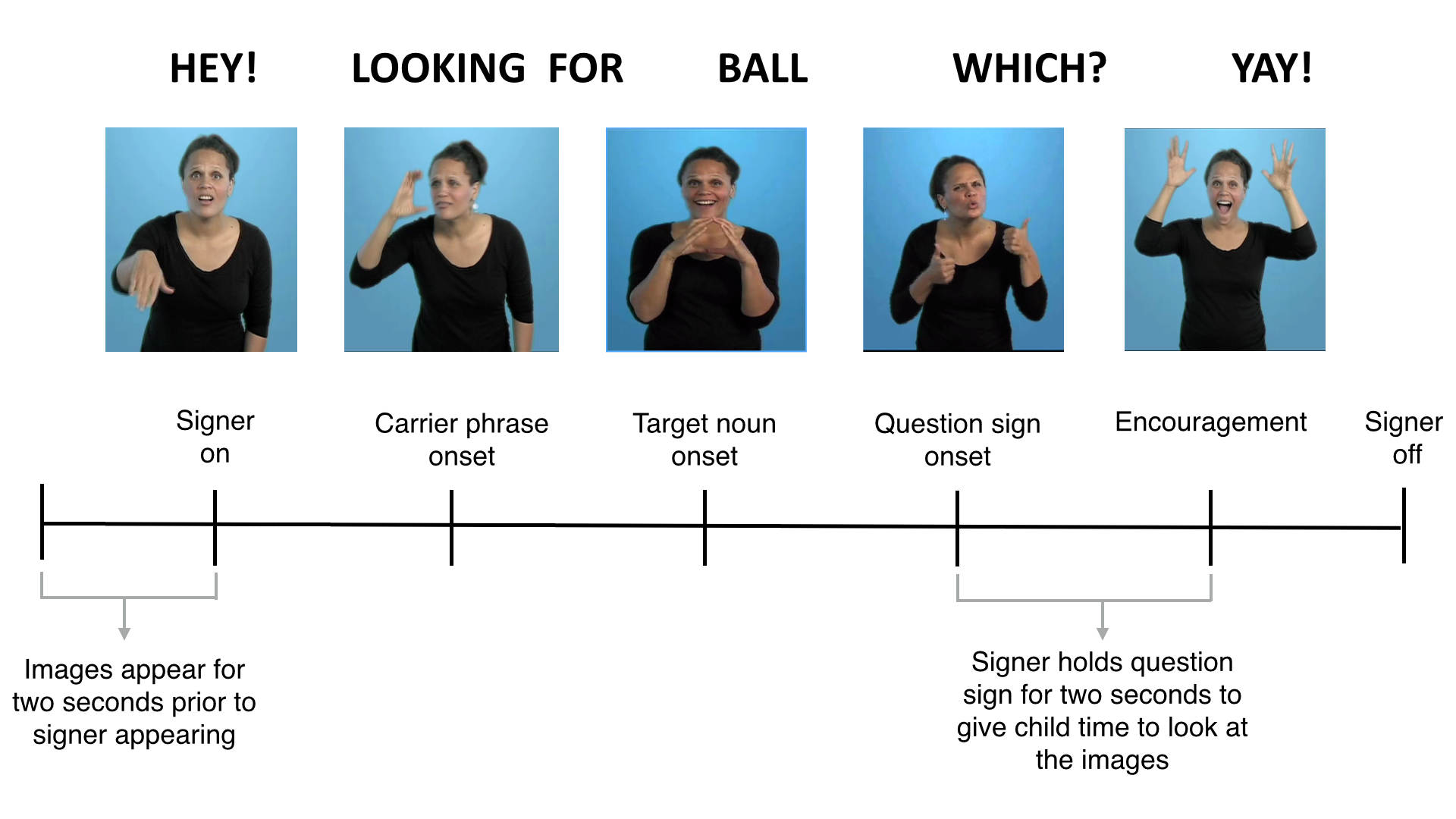
### Measures

*Parent report of 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 culturally and linguistically appropriate for children learning ASL. Parents completed the checklist during the visit, and 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 (LWL) (Fernald et al., 2008) adapted for ASL learners, which we call the Visual Language Processing (VLP) task. 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.

### Apparatus

In the VLP task, stimuli are presented on a 27” monitor using a Macbook Pro laptop. A video of the child’s gaze is recorded using a digital camcorder set up behind the monitor. To minimize visual distractions during the testing, the child sat on a caregiver’s lap inside of a portable 5’ by 5’ tent with opaque walls. The tent reduced the potential for visual distractions to occur during the task.

**Figure 1:** Timeline of a trial on the VLP task.

### Trial Structure

Figure 1 shows the timeline of one trial in the VLP task. On each trial, the child saw two images of familiar objects on the screen for two seconds 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 one second, which gave them the opportunity to orient to the signer prior to the sentence onset. The target sentence was then presented, followed by a question and hold, which gave the child the opportunity to shift their attention to the target object. After the hold, the signer gave neutral, positive feedback to help maintain the children's focus throughout the task. Each trial lasted approximately seven seconds.

### Linguistic and visual stimuli

The linguistic stimuli were designed to be comparable to those used in previous research and to allow for generalization beyond characteristics of a specific signer and sentence structure. To accomplish this, we recorded ASL sentences 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?”

Both versions of the stimuli can be viewed at the project page for this experiment: <https://github.com/kemacdonald/SOL>. We analyzed responses for the two sentence frames separately and found no significant differences between the two. All of the analyses we report collapse across the two sentence structures.

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 shift children's attention away from the images on the screen to the signer in preparation for the upcoming sentence.

Four yoked pairs of eight target nouns (cat—bird, car—book, bear—doll, ball—shoe) were used. These nouns were selected such that they would be familiar to most children learning ASL at this age and have minimal phonological overlap. To prepare the stimuli, two female native ASL users recorded several tokens of each sentence in a child-directed register. These candidate stimuli were then 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. Side of target picture was counterbalanced across trials.

### Coding and reliability

Children’s gaze patterns were videotaped and later coded frame-by-frame, yielding a high-resolution record of eye movements aligned with target noun onset. . Prior to coding, all sessions were pre-screened for parental interference. A total of 25% of videos were re-coded to assess coder reliability -- agreement within a single frame averaged 98% on these reliability assessments.

### Calculating linguistic processing efficiency

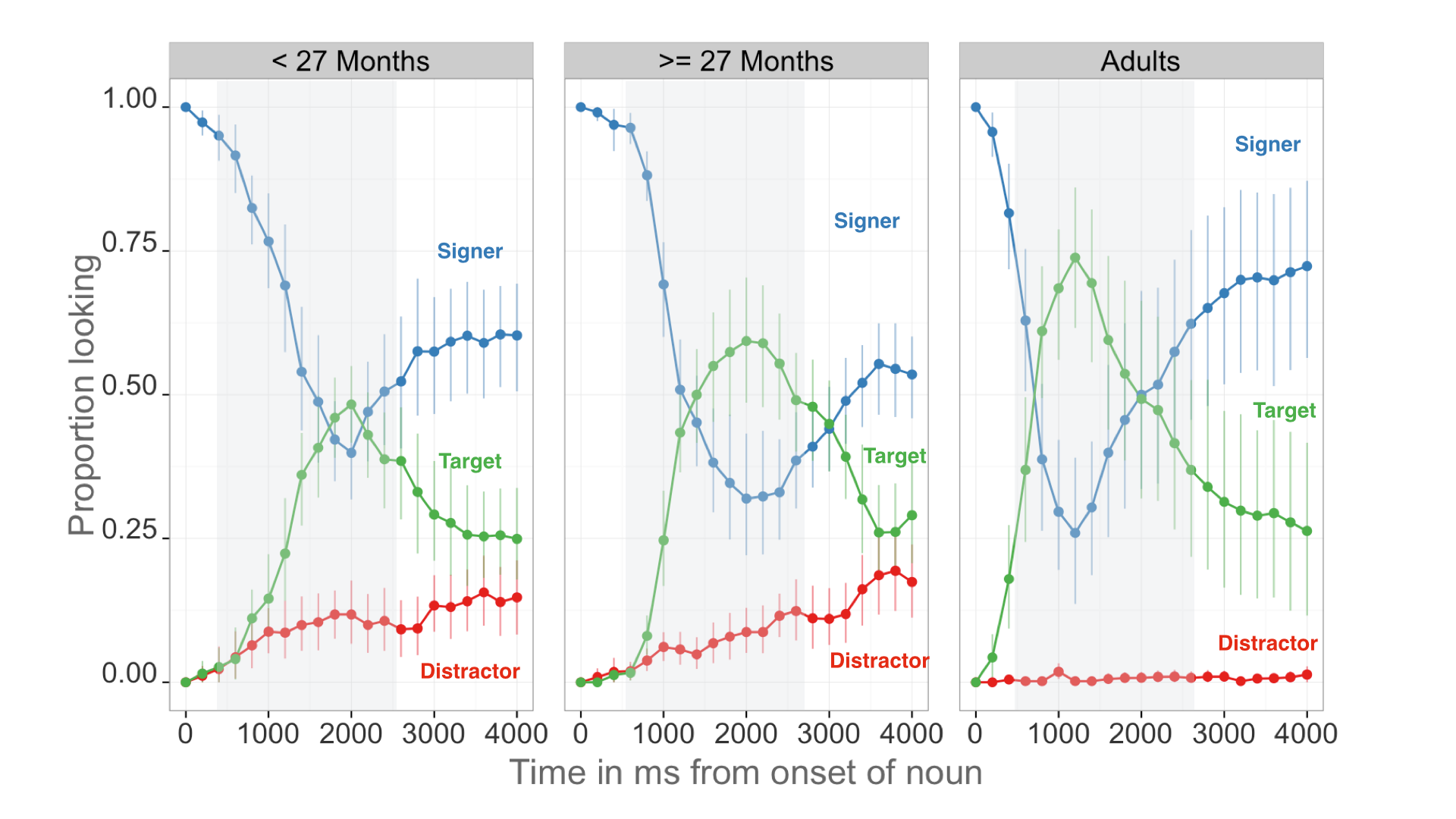
*Computing critical sign onset.* In the VLP task, computing accuracy and RT requires defining the appropriate response window, starting at the earliest moment when there is sufficient information to discriminate the pairs of pictures and to initiate a shift in gaze off of the central signer and to the target image. In studies of spoken language processing, critical word onset is typically identified using acoustic analysis software that measures the moment in the auditory signal when the target noun begins. In signed languages like ASL, phonological information is presented simultaneously in several parts of the visual signal (e.g., hands and face), so there is considerable co-articulation. It is therefore sometimes difficult to precisely determine the beginning of the target sign. This problem is somewhat simplified in the VLP task because the pictures are presented in yoked pairs, and so on each trial, critical sign onset is always determined in reference to particular distracter sign/picture. Prior work with adults has used different definitions of sign onset. For example, in the gating literature, sign onset has been defined as the moment when the signer’s hand(s): (a) appeared on the screen (Grosjean, 1981), (b) entered signing space (Emmorey & Corina, 1990), or (c) left the resting position (Arendsen, Van Doorn, & de Ridder, 2009). In one of the few tasks that used complete ASL sentences, Morford and Carlsen (2011) defined sign onset as “the frame in which the dominant signing hand moved out of the trajectory of the preceding sign and began the movement trajectory of the target sign.”

Here, we decided to take an empirical approach to defining critical sign onset. To get an initial judgment, the first and second authors, both fluent signers, viewed each stimulus sentence and achieved a consensus regarding target noun onset. Then, to validate these judgments, we conducted an additional study in which fluent adult signers unfamiliar with the stimuli (n = 10) watched videos of each signed sentence and made forced choice decisions indicating which of two images was signed in the video. For each sign, we selected six tokens, with each token containing a different amount of the target sign. These tokens were chosen to be plus or minus three frames of video from the noun onsets defined by the initial consensus judgments. Participants saw a total of 168 trials (28 sentences x 6 tokens), with order randomized across participants. We derived the final noun onset values for each stimulus sentence based on the earliest point when participants' judgments achieved 100% agreement. To our knowledge this is the first study to use empirically-derived sign onsets, but we felt it was critical to maximize the precision in our processing measures.

*Reaction Time:* In the VLP task, four different types of responses are possible on a given trial: (1) signer-to-target shift, (2) signer-to-distractor shift, (3) signer-to-away shift, (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. Incorrect shifts (signer-to-distracter, signer to away, no-shift) were not included in the computation of mean RT. 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 (600-2500 ms), since these data are the most likely to be generated by underlying process of interest: children’s lexical access. In addition, 8% of trials were excluded because children never shifted off of the signer. Since children vary in the likelihood that they will generate a correct shift, mean RTs are based on different numbers of trials across participants (M = 12.7 trials, range = 3—25).

Moreover, we excluded participants from the RT analysis if their first shift behavior was determined to be the result of guessing, since this suggests that RTs are not a meaningful measure of language processing skill for these children. To quantify the probability that participants' initial shifts were the result of guessing behavior, we used a Bayesian latent mixture model implemented in JAGS (Plummer, 2003). In this model, we assume that the observed data are generated by two different processes (guessing and knowledge) that have different 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 is a latent variable that we inferred based on their proportion of correct shifts to the target picture relative to the overall proportion of correct shifts across all participants (see Lee & Wagenmakers (2013) for a detailed discussion of this modeling approach). Five children were excluded because more than 50% of their posterior mass indicated a guessing strategy with relatively little uncertainty -- posterior probabilities of 0.99, 0.78, 0.98, 0.98, and 0.94 respectively (mean First Shift Accuracy scores for these participants were: 0.54, 0.58, 0.42, 0.29, 0.33).[[3]](#footnote-3) Thus, only signer-to-target shifts for 24 children were included in the RT analyses.

*Accuracy:* Correct looking is a function of the child’s tendency to shift quickly away from the central signer to the target picture in response to the target sign, and also to remain fixated on the target picture. To determine the degree to which participants fixated the appropriate 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, the

**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 (600-2500ms), and the error bars represent +/- 95% CI computed by non-parametric bootstrap.*

distracter picture, or the signer from 600 to 2500 ms from target noun onset. All children and response types were included in the computation of Accuracy, nevertheless, the number of trials contributing to the analysis varied across participants (M = 19.1, range = 6 to 35).

*First Shift Accuracy:* First Shift Accuracy was defined as the number of signer-to-target (correct) initial shifts divided by the number of signer-to-distractor (incorrect) shifts. All participants and trial types contributed to First Shift Accuracy analyses (M = 16.5, ranges = 5 to 30).

# Results

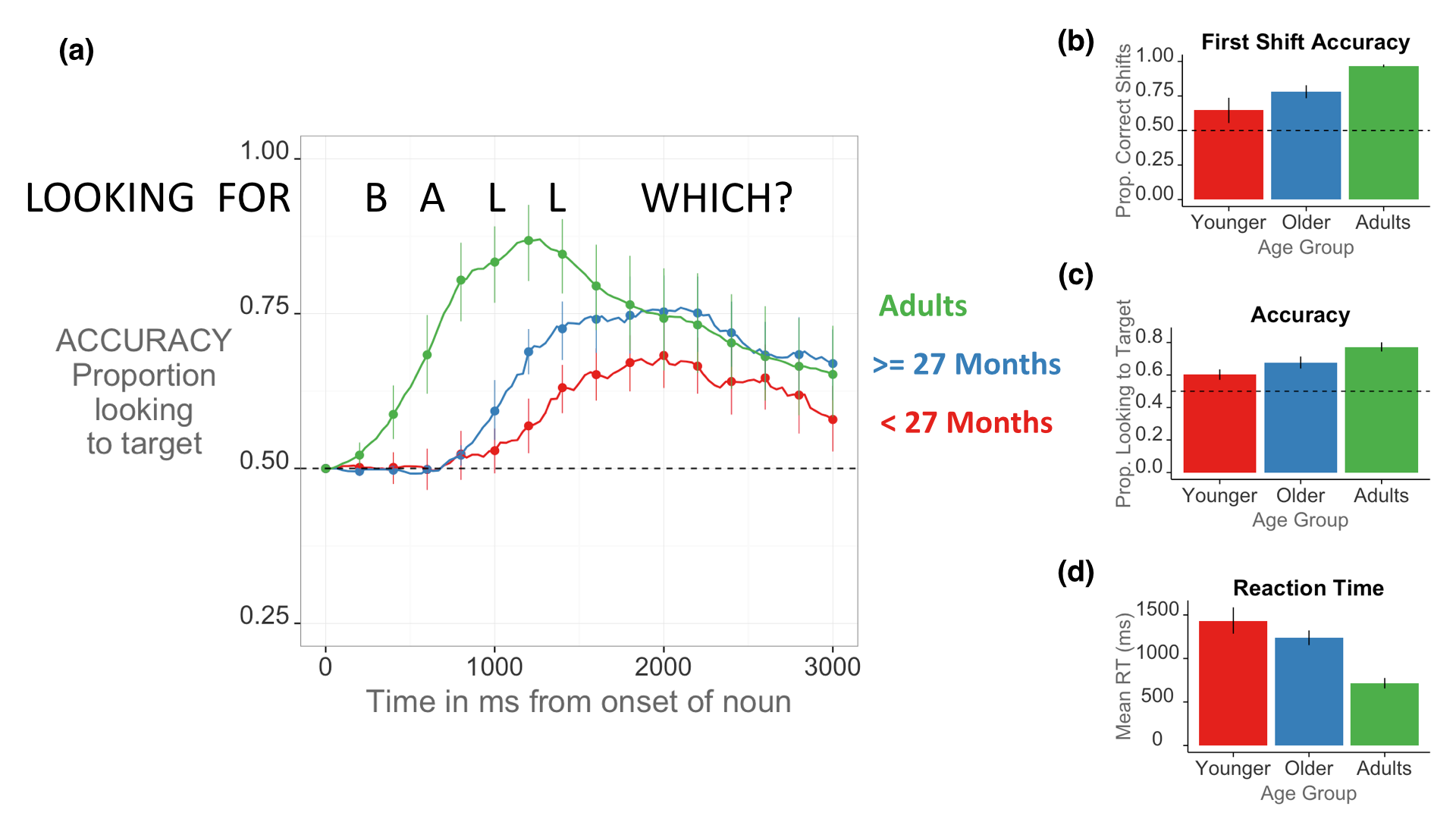
First, we present an overview of looking behavior on the VLP task. Then we show that children become faster and more accurate at comprehending familiar signs as they get older and make progress towards adult levels of language fluency. We use multiple regression analyses to model the links between children’s real-time ASL processing skills and both age and productive ASL vocabulary. Finally, we present a comparison of deaf and hearing ASL learners’ performance on the VLP task.

### Overview of ASL processing

Figure 2 provides an overview of the time-course of looking behavior in the VLP task. The three curves show changes in the mean proportion of trials on which participants in each age group fixated to the signer, the target image, or the distractor image at every 33 ms interval as the sign unfolded. At the onset of the target sign, all participants are looking at the signer. As the sign unfolded, proportion looking to the signer decreases rapidly as participants shift their gaze to either the target or the distractor image. Proportion looking to the target increases sooner in the sentence and reaches a higher asymptote compared to proportion looking to the distractor for all age groups, providing evidence that even the youngest children were capable of doing the task. After looking to the target image, participants’ tended to shift their gaze back to the signer, reflected by the increase in the blue curve. But, children also showed a small increase in looking to the distractor image towards the end of the trial.

### Measures of ASL processing efficiency

Figure 3 provides a graphical overview of all language-processing measures, and Table 1 shows summary statistics. Panel A shows correct orienting to the referent (Accuracy) in response to the target sign within the analysis window. At the onset of the target sign, mean proportion looking to target is 0.5, meaning that participants were fixated on the signer. As the sign unfolded, proportion looking to the target increased. However, children were slower to respond and less accurate, maintaining their gaze on the center signer longer and reaching a lower asymptote than did the adults. The youngest children less likely to initiate center-to-target shifts (t(27) = -2.51, d = 0.93), were less accurate (t(27) = -2.96, d = 1.1), and had longer mean latencies to orient to the target image (t(22) = 2.05, d = 0.57).

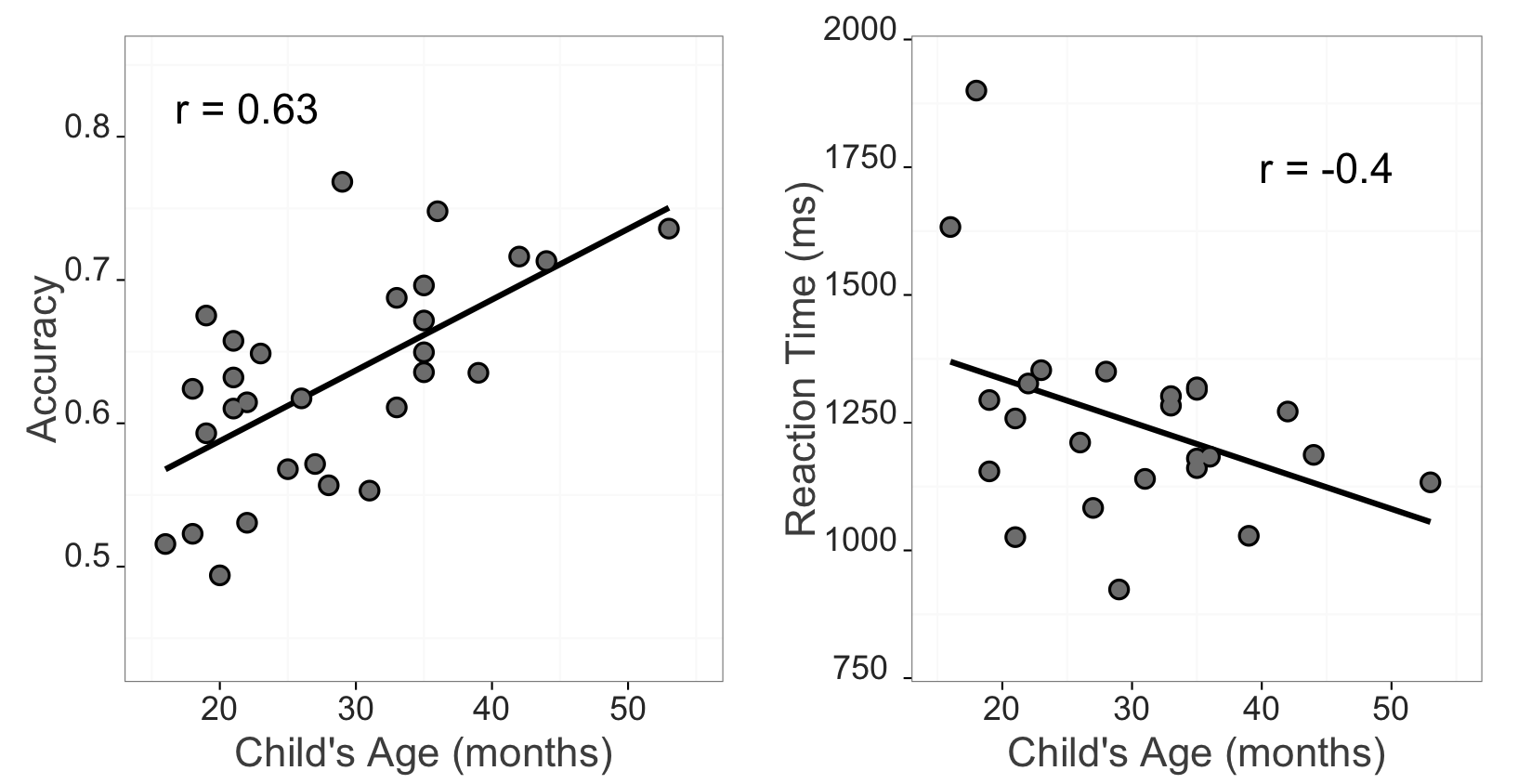
**Figure 3.** Panel A shows the time-course of participants' Accuracy for younger children, older children, and adults. Curves show changes over time in the mean proportion looking to the correct picture, measured in ms from noun onset; Panel B shows mean First Shift Accuracy; Panel C shows mean Accuracy. And Panel C shows mean RT. Error bars represent +/- 95% CI computed by non-parametric bootstrap.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Age Group** | **Accuracy** | **First Shift Accuracy** | **RT** | |
| < 27 Months | 0.59 (0.06) | 0.66 (0.20) |  | 1304 (255) | |
| >= 27 Months | 0.66 (0.07) | 0.80 (0.09) |  | 1191 (120) | |
| Adults | 0.80 (0.10) | 0.97 (0.03) |  | 699 (170) | |

**Table 1.** Means and standard deviations for all language processing measures for each age group.

### Links between processing efficiency and age

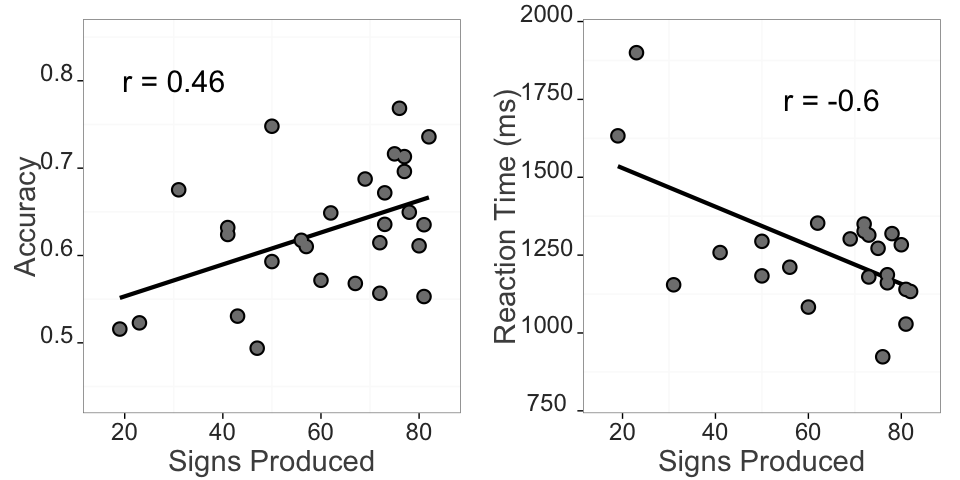
Mean accuracy scores, computed over the 600–2500 ms window from noun onset, were examined as a function of age. Accuracy was strongly correlated with age ((27) = 0.63), indicating that older ASL learners were more reliable than younger children in fixating the target picture. Mean reaction times were negatively correlated with age ((22) = -0.4), indicating that older children were faster to shift to the target picture than younger ones. Mean reaction times were also negatively correlated with mean accuracy scores ((22) = -0.6) such that those children who were faster to shift to the target were also more likely to stick on the target image throughout a greater proportion of the analysis window. Figure 4 shows these relations.

***Figure 4:*** *Scatterplot of the relation between scores on the VLP measures and children’s age. Left Panel shows the positive relationship between accuracy and age. Right panel shows the negative relationship between RT and age.*

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, progressing towards adult levels of language fluency.

### Links between processing efficiency and vocabulary

Figure 5 shows the relationships between both VLP processing measures and children's productive ASL vocabulary. Mean accuracy was positively related to vocabulary size ((27) = 0.46) such that children with higher accuracy scores also had larger productive vocabularies. Mean reaction times were negatively correlated with vocabulary ((22) = -0.6) indicating that children who were faster to recognize ASL signs were those children with larger sign vocabularies.

**Figure 5:** Relation between VLP measures and productive ASL vocabulary. Left panel shows the positive relationship between accuracy and vocabulary. Right panel shows the negative relationship between RT and vocabulary.

It is important to point out that age and vocabulary were strongly intercorrelated in our sample ((27) = 0.74). So in order to quantify the effect of each factor on children's online language processing skills, we fit separate multiple regression models to children's Accuracy and RT data. Table 2 shows the Accuracy model. Age (months) and vocabulary (number of signs produced) were coded as continuous predictors, and accuracy was coded as the mean proportion correct looking to the target for each child. Together, age and vocabulary accounted for approximately 42.2% of the variance in accuracy ((2, 25) = 9.1), however, only age contributed significant unique variance (-change: 18.%).Thus, the majority of the variation in accuracy was attributable to the shared variance between these two factors, yet some sources of individual differences in accuracy were attributable to age above and beyond vocabulary.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | Estimate | Std. Error | t.value | p.value |
| **Intercept** | 0.486 | 0.0399 | 12.2 | < .001 |
| **Vocabulary** | -0.000346 | 0.000939 | -0.368 | 0.716 |
| **Age** | 0.00567 | 0.00188 | 3.02 | < .001 |

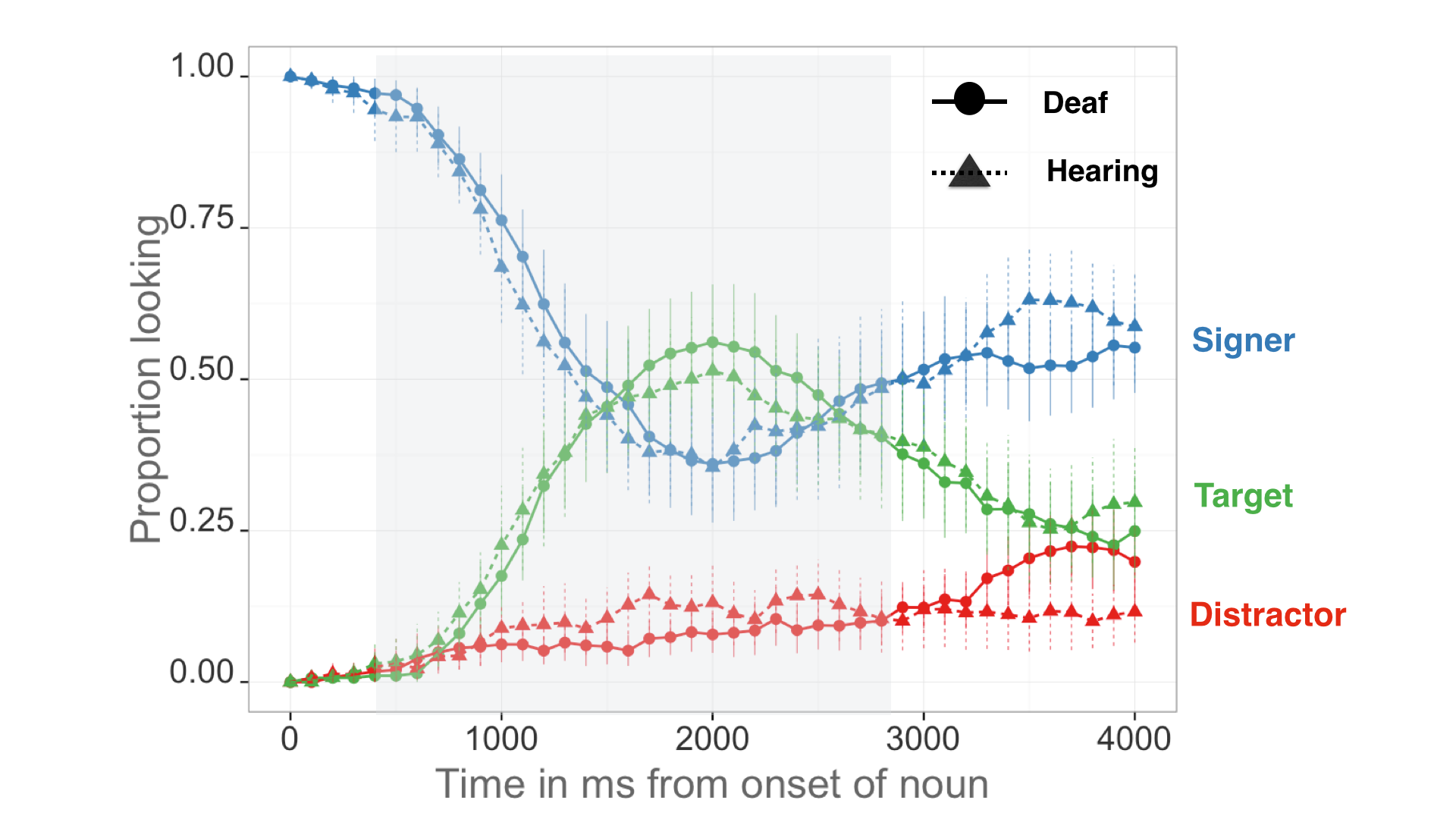
***Table 2:*** *Multiple regression model output for accuracy data. The model was specified as: Accuracy ~ Vocabulary + Age.*

Next, we modeled predictors of children's RT. Table 3 shows that age and vocabulary together accounted for approximately 36.5% of the variance in RT. However, in contrast to accuracy, only vocabulary remained a significant predictor after including age in the model (r2change = 16%). Thus, children who were faster to identify the target sign also had larger vocabularies, even after controlling for age.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | Estimate | Std. Error | t.value | p.value |
| **Intercept** | 1660 | 128 | 13 | < .001 |
| **Vocabulary** | -5.87 | 2.77 | -2.12 | 0.046 |
| **Age** | -0.901 | 5.68 | -0.159 | 0.876 |

***Table 3:*** *Multiple regression model output for RT data. The model was specified as: RT ~ Vocabulary + Age*

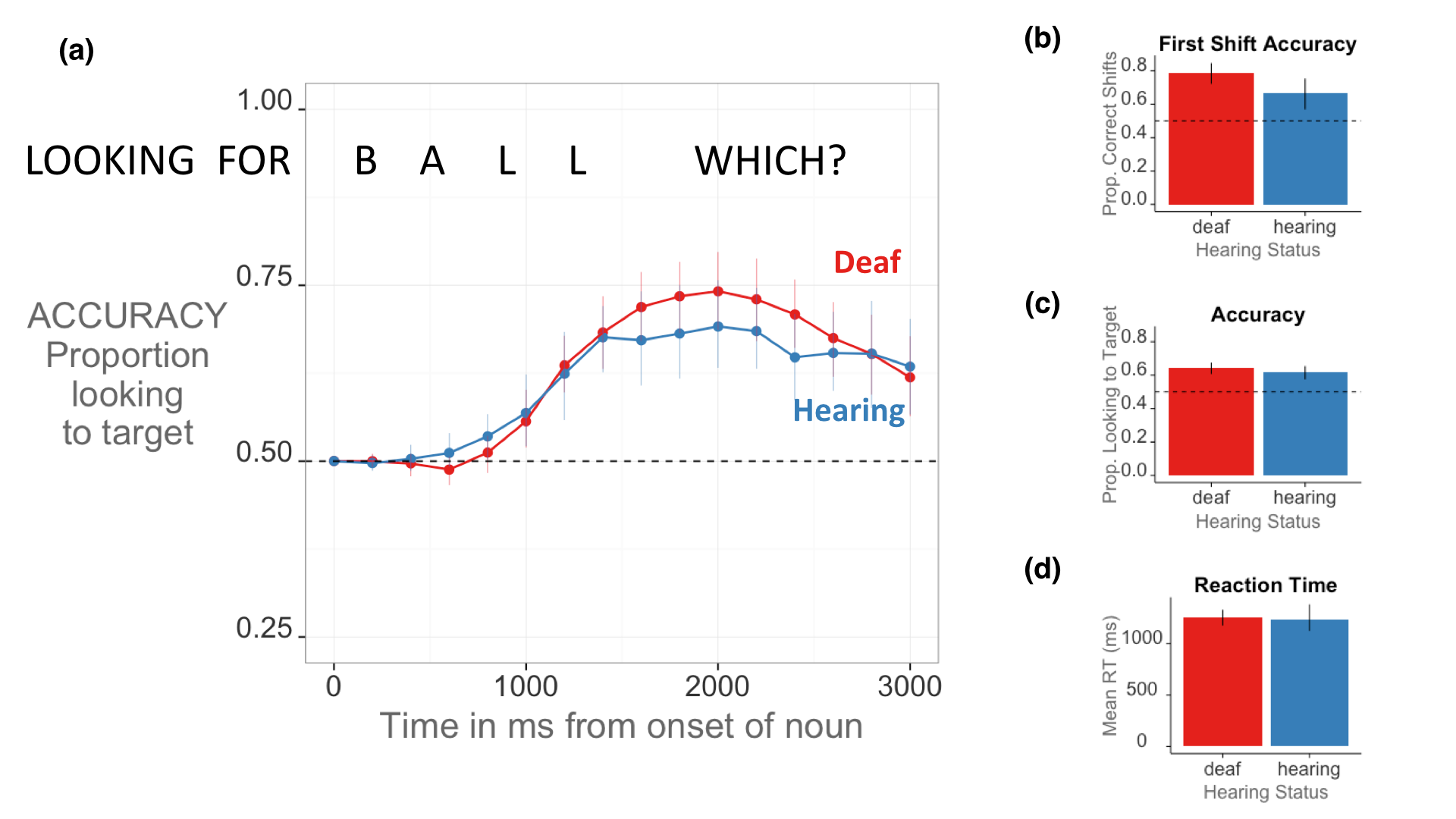
Taken together, these analyses indicate that children learning ASL were more accurate and efficient in identifying the referents of familiar signs as they got older and developed a larger expressive vocabulary. These findings are consistent with previous research with children learning English and Spanish (Fernald et al., 2006; Hurtado, Marchman, & Fernald, 2007).

*****Figure 6.*** *The time-course of looking behavior for deaf and hearing children. The curves show proportion looking to the signer (blue), the target image (green), and the distractor image (red). The shape and line type represent hearing status; the grey shaded region represents the analysis window (600-2500ms); error bars represent +/- 95% CI computed by non-parametric bootstrap.*

### Effects of hearing status

Figure 6 shows an overview of performance 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 remarkably 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.

Figure 7 shows performance across all VLP measures for both groups. We found no differences in Accuracy (t(27) = 0.93, ns.) or Reaction Time (t(22) = 0.16, ns.). But we did find a marginally significant difference in First Shift Accuracy (t(27) = 2.02, d = 0.78), with deaf children (M = 0.79) being more accurate with their initial shifts off the signer than hearing children (M = 0.66). Taken together, these analyses provide evidence that auditory experience does not qualitatively change visual language processing. However, there was some evidence that initial eye movements were more accurate for deaf children.



***Figure 7.*** *Panel A shows the time course of Accuracy for deaf and hearing children. The curves show changes over time in the mean proportion looking to the correct picture, measured in ms from noun onset. Panel B shows mean First Shift Accuracy. Panel C shows mean Accuracy. And Panel D shows mean Reaction Time. Error bars represent +/- 95% CI computed by non-parametric bootstrap.*

# Discussion

Establishing reference in real-time is a fundamental component of language learning. To link signs to objects, young ASL users must learn to resolve an apparent conflict between attending to the source of linguistic information and shifting their gaze to the surrounding visual scene. Moreover, they must learn to do this efficiently because if a child does not see a sign, or does not see the intended referent, then the information in that naming event is effectively unavailable. With this study, we aimed to develop and validate the first measures of young ASL learners’ real-time language comprehension skills and explore the links between these skills and age, vocabulary, and hearing status. There are four main findings from this work.

The first main finding was that, like children learning spoken language (Fernald et al., 1998), young ASL learners' show measurable age-related improvement in the efficiency with which they processed language. All of the target signs were familiar to children in this age range, yet older children more quickly and accurately identified the correct referents than younger children. This finding provides additional evidence that ASL acquisition in native, natural contexts follows a strikingly similar developmental path to 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 development. Prior work on the developmental trajectories of deaf children has relied on aspects of language production often because production is easier to see, making it easier to measure. However, it is a well-known phenomenon in language acquisition that comprehension tends to precede production (see Clark, 2009). By developing a fine-grained measure of real-time ASL comprehension, we have provided a tool that can measure children's language skills earlier in development than previously possible.

The second result was the discovery of a link between early ASL processing skills and children's productive ASL vocabularies, measured independently using parent report. ASL-learning children who knew more signs were also faster and more accurate in language processing than those who were lexically less advanced. However, the factors of age and vocabulary size were highly intercorrelated in this sample and the majority of the association between vocabulary and accuracy, but not RT, was attributable to variance that was shared between these two factors. Nevertheless, 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 was that native ASL learning deaf and hearing children show remarkably 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 similarities suggest that auditory experience does not qualitatively change visual language processing. We did, however, find some evidence that deaf children’s initial gaze shifts were more accurate compared to hearing children. It is important to point out that this result was based on exploratory analyses of these data, and it should be replicated before making strong claims about differences between deaf and hearing children’s ASL processing. With that being said, it is interesting to speculate about the source of the difference in first shift accuracy. Perhaps deaf children are more accurate because gaze shifts are more costly for this group. If vision is the only way to monitor the world, then deaf children might require more information or a higher level of confidence before they initiate a shift away from a signer. Ongoing studies in our laboratory are attempting to tease apart these interesting speed-accuracy tradeoffs.

Taken together, these age, vocabulary, and hearing status results lead us to fourth main finding: The VLP task captures meaningful individual variation in visual language processing. In adults, previous work using online measures with adult ASL users has revealed important aspects of the psycholinguistics of sentence processing. For example, Lieberman et al. (2014a) showed evidence of real-time activation of sublexical features of sign and that incremental semantic processing occurs during real-time sign comprehension. But we did not yet know whether these online measures would work with very young learners early in the development of sign language proficiency. Moreover, because gaze serves a variety of functions in ASL (e.g., process visual and linguistic information), using eye movements as an index of language comprehension skills might not have been possible. But, despite these modality-driven differences, the current study shows that the VLP task is able to measure subtle individual differences in real-time ASL processing skills, suggesting that this task may be very useful in further studies in the development of sign language. It is interesting to consider how the gaze shifts we measured would compare to the gaze shifting behavior required in naturalistic interactions that are thought to be critical for early lexical development.

### Limitations

This study has several limitations that are worth noting. First, while our sample is large relative to past work investigating early ASL development, it is still a small sample. These 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 (<https://github.com/kemacdonald/SOL>), with the hope that other researchers will benefit from what we have learned in this work.

Second, our sample included a broad age range of children. Recall that age explained the majority of individual variation in accuracy, but, interestingly, not in RT. Based on the strength of past evidence, it is likely that testing a narrower age range would allow us to see an independent effect of vocabulary size in both language processing measures. We need more evidence in order to best characterize the relationships between accuracy, RT, and vocabulary.

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 (e.g., Emmorey & Corina, 1990; Morford & Carlsen, 2011), our stimuli were full sentences signed in a child-directed manner, 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 contains a natural central fixation point – the signer – making it substantially different from previous work on 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 again show remarkable similarities with previous findings with learners of spoken languages in terms of age-related changes and links to other 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 from at least one deaf caregiver. However, we might anticipate that the development of online language processing might look different in children who are late learners or who have much more heterogeneous and inconsistent language exposure than the users of ASL who were tested in the current work. An important next step is to explore the relations between the development of language processing skills and children's experience with signed languages during interactions with caregivers, the norm for most children born deaf and learning a visual language. It is well-known that children's efficiency of online processing is tightly 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 eye movements during real-time ASL sentence processing provide a valid measure of age-related changes in young children's visual language comprehension skills. These findings contribute to the now significant body of literature highlighting the parallels between signed and spoken language development when children are exposed to native sign input. We hope that the development of the VLP task will provide a useful method for researchers and educators, providing a way to track developmental trajectories of early language learning in native learners of signed languages, like ASL.

# References (56)

Altmann, G. T., & Kamide, Y. (1999). Incremental interpretation at verbs: Restricting the domain of subsequent reference. *Cognition*, *73*(3), 247–264.

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.

Bahan, B., & Supalla, S. (1995). Line segmentation and narrative structure: A study of eyegaze behavior in American Sign Language. *Language, Gesture, and Space*, 171–191.

Baker, C., & Padden, C. (1978). Focusing on the nonmanual components of American Sign Language. *Understanding Language Through Sign Language Research*, 27–57.

Baldwin, D. A. (1995). Understanding the link between joint attention and language.

Bellugi, U. (1979). *The signs of language*. Harvard University Press.

Bosworth, R. G., & Dobkins, K. R. (2002). The effects of spatial attention on motion processing in deaf signers, hearing signers, and hearing nonsigners.*Brain and Cognition*, *49*(1), 152-169.

Boudreault, P., & Mayberry, R. I. (2006). Grammatical processing in American Sign Language: Age of first-language acquisition effects in relation to syntactic structure. *Language and Cognitive Processes*, *21*(5), 608–635.

Brooks, R., & Meltzoff, A. N. (2008). Infant gaze following and pointing predict accelerated vocabulary growth through two years of age: A longitudinal, growth curve modeling study. *Journal of Child Language*, *35*(01), 207–220.

Carreiras, M. (2010). Sign language processing. *Language and Linguistics Compass*, *4*(7), 430–444.

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.

Dahan, D., & Tanenhaus, M. K. (2004). Continuous mapping from sound to meaning in spoken-language comprehension: Immediate effects of verb-based thematic constraints. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *30*(2), 498.

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

Emmorey, K., Kosslyn, S. M., & Bellugi, U. (1993). Visual imagery and visual-spatial language: Enhanced imagery abilities in deaf and hearing ASL signers. *Cognition*, *46*(2), 139-181.

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., Weinbergy, A., & McRoberts, G. W. (1998). Rapid gains in speed of verbal processing by infants in the 2nd year. *Psychological Science*, *9*(3), 228–231.

Fernald, A., Swingley, D., & Pinto, J. P. (2001). When half a word is enough: Infants can recognize spoken words using partial phonetic information. *Child Development*, *72*(4), 1003–1015.

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.

Hildebrandt, U., & Corina, D. (2002). Phonological similarity in American Sign Language. *Language and Cognitive Processes*, *17*(6), 593–612.

Hurtado, N., Marchman, V. A., & Fernald, A. (2007). Spoken word recognition by lLatino 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*.

Lew-Williams, C., & Fernald, A. (2007). Young children learning Spanish make rapid use of grammatical gender in spoken word recognition. *Psychological Science*, *18*(3), 193–198.

Lieberman, A. M., Borovsky, A., Hatrak, M., & Mayberry, R. I. (2014a). 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. (2014b). 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.

Lillo-Martin, D. (2000). Early and late in language acquisition: Aspects of the syntax and acquisition of wh-questions in American Sign Language. *The Signs of Language Revisited: An Anthology to Honor Ursula Bellugi and Edward Klima*, 71–90.

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., & Eichen, E. B. (1991). The long-lasting advantage of learning sign language in childhood: Another look at the critical period for language acquisition. *Journal of Memory and Language*, *30*(4), 486–512.

Mayberry, R. I., & Fischer, S. D. (1989). Looking through phonological shape to lexical meaning: The bottleneck of non-native sign language processing. *Memory & Cognition*, *17*(6), 740–754.

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

McClelland, G., Lynch, J. G., Irwin, J. R., Spiller, S. A., & Fitzsimons, G. J. (2015). Median splits, type II errors, and false positive consumer psychology: Don’t fight the power. *Type II Errors, and False Positive Consumer Psychology: Don’t Fight the Power (March 10, 2015)*.

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.

Neidle, C., MacLaughlin, D., Lee, R. G., Bahan, B., & Kegl, J. (1998). Wh-questions in ASL: A case for rightward movement. *American Sign Language Linguistic Research Project*, *6*.

Newport, E. L. (1990). Maturational constraints on language learning. *Cognitive Science*, *14*(1), 11–28.

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

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.

Quine, W. (1960). *Word and object*. MIT press.

Ratcliff, R. (1993). Methods for dealing with reaction time outliers. *Psychological bulletin*, *114*(3), 510.

Snedeker, J., & Trueswell, J. C. (2004). The developing constraints on parsing decisions: The role of lexical-biases and referential scenes in child and adult sentence processing. *Cognitive Psychology*, *49*(3), 238–299.

Tanenhaus, M. K., Spivey-Knowlton, M. J., Eberhard, K. M., & Sedivy, J. C. (1995). Integration of visual and linguistic information in spoken language comprehension. *Science*, *268*(5217), 1632–1634.

Thompson, R., Vinson, D., Fox, N., & Vigliocco, G. (2013). Is lexical access driven by temporal order or perceptual salience? Evidence from British sign language. In *Proceedings of the 35th annual meeting of the cognitive science society* (pp. 1450–1455).

Waxman, 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., & Fernald, A. (2007). Increasing flexibility in children’s online processing of grammatical and nonce determiners in fluent speech. *Language Learning and Development*, *3*(3), 199–231.

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. See McClelland, Lynch, Irwin, Spiller, & Fitzsimons (2015) for a discussion of the potential for increasing Type I and Type II errors through the loss of power caused by using median splits. [↑](#footnote-ref-1)
2. See Neidle, MacLaughlin, Lee, Bahan, & Kegl (1998) for a detailed discussion of the acceptability of these two question structures. [↑](#footnote-ref-2)
3. See the supplementary materials for more details about the Bayesian guessing model. [↑](#footnote-ref-3)