Exploring socioeconomic differences in syntactic development through the lens of real-time processing

Yi Ting Huang,1 Kathryn Leech,2,3 and Meredith L. Rowe3

1. University of Maryland College Park, Department of Hearing and Speech Sciences
2. University of Maryland College Park, Department of Human Development and Quantitative Methodology
3. Harvard University, Graduate School of Education

Corresponding Author:

Yi Ting Huang
Department of Hearing and Speech Sciences
University of Maryland College Park
College Park, MD 20742
ythuang1@umd.edu
Abstract

Differences in caregiver input across socioeconomic status (SES) predict syntactic development, but the mechanisms are not well understood. Input effects may reflect the exposure needed to acquire syntactic representations during learning (e.g., does the child have the relevant structures for passive sentences?) or access this knowledge during communication (e.g., can she use the past participle to infer the meaning of passives?). Using an eye-tracking and act-out paradigm, the current study distinguishes these mechanisms by comparing the interpretation of actives and passives in 3- to 7-year-olds (n = 129) from varying SES backgrounds. During the presentation of spoken sentences, fixations revealed robust disambiguation of constructions by children from higher-SES backgrounds, but less sensitivity by lower-SES counterparts. After sentence presentation, decreased sensitivity generated interpretive challenges and SES-related differences for passives requiring syntactic revision (“The seal is quickly eaten by it”). Critically, no differences were found when revision was not needed (“It is quickly eaten by the seal”), suggesting that all children shared a basic ability to acquire passive structures. These results suggest that SES-related differences are present in real-time syntactic processing and impact the accuracy of utterance interpretation.

Keywords: Socioeconomic status (SES); language processing; syntactic development; passive construction
1. Introduction

Striking differences in vocabulary development have been found in language acquisition across socioeconomic status (SES) (Arriaga, Fenson, Cronan, & Pethick, 1998; Hart & Risley, 1995; Hoff, 2003; Weisleder & Fernald, 2013). Notably, these effects are also present in syntactic development (Hoff-Ginsberg, 1986; Huttenlocher, Vasilyeva, Cymerman, & Levine, 2002; Vasilyeva, Waterfall, & Huttenlocher, 2008; Huttenlocher, Waterfall, Vasilyeva, Vevea, & Hedges, 2010; Huttenlocher, Vasilyeva, Waterfall, Vevea, & Hedges, 2007; Dollaghan et al., 1999; Morisset, Barnard, Greenberg, Booth, & Spieker, 1990), an area that is traditionally argued to be resilient to variation in learning environments (Newport, Gleitman, & Gleitman, 1977; Borer & Wexler, 1992). Relative to lower-SES counterparts, children from higher-SES backgrounds, on average, produce more complex utterances (e.g., number of clauses, words per sentence) and diverse constructions (e.g., number of structural relationships) (Huttenlocher et al., 2010; Vasilyeva et al., 2008). These distinctions are mirrored in the communicative input to children from varying SES backgrounds (Hart & Risley, 1995; Hoff, 2003; Huttenlocher et al., 2007; Rowe, 2012; Cartmill et al., 2013). Compared to lower-SES counterparts, caregivers from higher-SES backgrounds, on average, produce more complex syntactic structures such as wh-questions, relative clauses, and raising adjectives (Huttenlocher et al., 2010).

Yet, far less is known about why relationships between language outcomes and caregiver input emerge in the first place or what aspects of development they reflect. One possibility is that SES-related differences reflect variation in learning. If specific language experiences (i.e., input quantity or quality) are required to acquire syntactic representations, then children may simply fail to learn constructions that are not frequently encountered. A second possibility is that SES-related effects are far more targeted. While children may acquire syntactic knowledge with minimal experience, input properties may facilitate access to this knowledge during real-time
comprehension. If so, then SES-related differences may be isolated to situations where efficient access to previously acquired representations is necessary for the interpretation of an utterance. However, when utterance interpretation does not depend on efficient access to representations, then SES-related differences may be minimal.

To distinguish between effects of language experience during learning (e.g., does the child have syntactic representations?) versus communication (e.g., can she access them when she hears utterances?), the current study examines SES-related differences in the comprehension of a low-frequency construction: the English *be*-passives. In the remainder of the Introduction, we will flesh out two perspectives on the role of caregiver input during syntactic development and examine their predictions for the scope of SES-related effects. We will then briefly consider why findings from prior research fail to distinguish between these hypotheses and discuss how the current study will tackle these limitations.

1.1 Two perspectives on input effects during language development

Accounts of SES-related effects on syntactic development often focus on how language experience impacts the acquisition of linguistic representations at the point of learning. As such, there is an underlying assumption that variable outcomes reflect differences in *forming* syntactic structures via frequency-driven associations between caregiver input and utterance meaning (Huttenlocher et al., 2002; Huttenlocher et al., 2007; Huttenlocher et al., 2010). These accounts share similarities to influential theories of acquisition including social-interactionist (Snow, 1989; Bruner, 1983) and usage-based approaches (Tomasello, 2000; Ambridge, Kidd, Rowland, & Theakston, 2015). They also provide an intuitive explanation for why SES-related effects are present in syntactic development. Since learning is predicated on adequate language experience, it is unsurprising that children from lower-SES backgrounds (who encounter less quantity and quality of input) lag behind their higher-SES counterparts (who encounter more), on average.
However, far less work has considered the impacts of language experience not at the point of acquiring representations, but when accessing this knowledge during communication. Even among highly experienced language users like adults, variation in input statistics influences real-time syntactic interpretation (MacDonald, Pearlmutter, & Seidenberg, 1994; Trueswell & Tanenhaus, 1994; MacDonald, 2013; Wells, Christiansen, Race, Acheson, & MacDonald, 2009). These effects may have even greater consequences for less experienced language users such as children, who often fail to revise incremental misinterpretations, even after hearing conflicting syntactic cues in the utterance (Trueswell et al., 1999; Hurewitz, Brown-Schmidt, Thorpe, Gleitman, & Trueswell, 2000; Weighall, 2008; Choi & Trueswell, 2010; Huang, Zheng, Meng, & Snedeker, 2013; Omaki, Davidson White, Goro, Lidz, & Phillips, 2014). Thus, rather than enabling the initial formation of syntactic representations, this account suggests instead that language experience may promote real-time retrieval of these representations, particularly in contexts where efficient access is necessary for accurate utterance interpretation.

Indeed, research by Fernald and colleagues points to such a link in the area of vocabulary development (Fernald, Marchman, & Weisleder, 2013; Hurtado, Marchman, & Fernald, 2008; Weisleder & Fernald, 2013). On average, 18-month-olds from lower SES families were slower to recognize highly familiar words in spoken sentences compared to their peers from higher SES families (Fernald et al., 2013). Moreover, individual variation in the speed of lexical processing predicted vocabulary size six months later, suggesting that real-time comprehension may mediate relationships between language experience and vocabulary development. Nevertheless, the studies to date have focused on word recognition in simple and frequent syntactic contexts (e.g., “Where’s the dog?”). Thus, it remains unknown how these effects translate to development at later ages. Do SES-related delays also emerge in syntactic processing? If so, how do they impact children’s interpretation of spoken utterances?
1.2 Why comprehension of the passive construction may be informative

One challenge to addressing these questions is the traditional reliance on aggregated measures of language performance (e.g., mean length utterance, number of clauses, standardized assessments). These tools provide excellent summaries of the range of SES-related effects, but they can also obscure their underlying causes. In particular, these approaches fail to distinguish between whether SES-related differences in language comprehension reflect variation in the acquisition of syntactic representations (i.e., properties of caregiver input enables some but not all children to learn syntactic structures) versus real-time retrieval during communication (i.e., all children have knowledge of structures, but caregiver input enables some to access this knowledge more efficiently). Thus, to isolate the mechanisms underlying SES-related differences during syntactic development, it is necessary to adopt finer-grained measures of performance.

To this end, the current study focuses on children’s comprehension of a well-studied test case in language acquisition: the active-passive alternation. Both constructions express the basic relationship of who did what to whom. In actives like (1a), the first noun phrase (NP1) maps onto the agent (“the seal” = PREDATOR) while the second noun phrase (NP2) maps onto the theme (“the fish” = PREY). In passives like (1b), this order is reversed: NP1 is now the theme (“the seal” = PREY) while NP2 is the agent (“the shark” = PREDATOR). It is well documented that children readily comprehend actives, but often generate errors with passives (Maratsos, Fox, Becker, & Chalkley, 1985; Gordon & Chafetz, 1990; Harris & Flora, 1982; Huang et al., 2013; Stromswold, Eisenband, Norland, & Ratzan, 2002; Sudhalter & Braine, 1985; Messenger, Branigan, & McLean, 2012). This asymmetry has inspired several theories of syntactic development (see Huang et al., 2013 for a review). For our present purposes, we focus on two prominent accounts and consider their predictions for SES-related differences.

(1)  a. Active: The seal is quickly eating the fish.
b. Passive: The seal is quickly eaten by the shark.

First, many have argued that children’s difficulties with passives reflect their lack of experience with the construction (Harris & Flora, 1982; Gordon & Chafetz, 1990; Demuth, 1989; Brooks & Tomasello, 1999). Passives are far less frequent than actives in the input. Stromswold and colleagues (2002) found that full passives accounted for less than 0.2% of adult utterances to children (see also calculations by Maratsos et al., 1985 and Gordon & Chafetz, 1990). Moreover, earlier proficiency is often found in languages where passives are more frequent, e.g., Inuktitut (Allen & Crago, 1996), K’iche’ Mayan (Pye & Poz, 1988), Sesotho (Demuth, 1989, 1990). It is plausible that SES-related effects on syntactic development reflect frequency-driven differences in acquisition that emerge within a single language. While we know of no study to date that has examined SES-related differences in input to passives specifically, related distinctions are well-documented with other complex structures (Huttenlocher et al., 2010). Thus, to the extent that low-frequency passives may be even less frequent for children from lower-SES backgrounds, it is possible that these children would be less likely to acquire this construction.

Nevertheless, errors with passives are also found among adults (Ferreira, 2003; Huang et al., 2013; MacWhinney, Bates, & Kliegl, 1984), suggesting that their underlying cause may not necessarily be a lack of knowledge. Instead, comprehension errors may reflect properties of the construction that pose challenges for real-time processing (Turner & Rommetveit, 1967; Bever, 1970; Huang et al., 2013). Since passives cannot be distinguished from actives until after the onset of verb morphology (see (1)), children may initially misanalyze NP1s as agents and fail to revise this interpretation, even after hearing conflicting linguistic cues (e.g., past participle, by-phrase). Challenges with syntactic revision are well documented in school-aged children (Trueswell et al., 1999; Hurewitz et al., 2000; Weighall, 2008; Choi & Trueswell, 2010; Huang et al., 2013; Omaki et al., 2014). Moreover, it is striking that SES-related effects of input are
often found in structures that require interpretation of arguments displaced from their canonical positions (Huttenlocher et al., 2010). Thus, if efficient syntactic revision depends, in part, on encountering structures that require reanalysis, then SES-related differences may specifically arise when comprehending passives, which often require revision.

1.3 The current study

The following experiment distinguishes these accounts by comparing the comprehension of active and passive sentences in 3- to 7-year-old children, from varying SES backgrounds. This age range lies at the intersection of three relevant literatures: (1) SES-related effects on syntactic development (Hoff-Ginsberg, 1986; Huttenlocher et al., 2002, 2007, 2010; Vasilyeva et al., 2012), (2) developmental difficulties with passives (Maratsos et al., 1985; Sudhalter & Braine, 1985; Gordon & Chafetz, 1990; Stromswold et al., 2002; Messenger et al., 2012; Huang et al., 2013), and (3) developmental difficulties with syntactic revision (Trueswell et al., 1999; Hurewitz et al, 2000; Weighall, 2008; Choi & Trueswell, 2010; Huang et al., 2013; Omaki et al., 2014). The current study recruits a language-processing task developed by Huang and colleagues (2013), which combines an eye-tracking paradigm (to assess on-line sensitivity to syntactic cues) with an act-out task (to assess the accuracy of utterance interpretation).

Children were presented with a series of active and passive sentences (Table 1), paired with thematically related objects: Expressed item (SEAL), likely agent (SHARK), and likely theme (FISH). Following verb morphology (e.g., “eating” vs. “eaten”), preferences for the likely agent or theme provide an implicit measure of role assignment. In the Expressed NP1 condition, expressed nouns in actives are agents (e.g., “the seal”), thus pronoun NP2s are likely themes (e.g., “it” = FISH). In contrast, when expressed nouns in passives are themes, pronoun NP2s are likely agents (e.g., “it” = SHARK). When presented with Mandarin versions of these sentences (Huang et al., 2013), 5-year-olds from higher-SES backgrounds were less accurate with passives
compared to actives (52% vs. 74%). This pattern is consistent with a frequency account, which argues that children lack structures for low-frequency constructions. It is also consistent with a processing account, which argues that inefficient retrieval of low-frequency structures hinders revision of an agent-first bias. Both accounts predict that these challenges may be magnified in children from lower-SES backgrounds, who may have even less experience with passives.

Table 1. Sample sentences in the four critical conditions of the language-processing task. Each sentence was paired with a three-object set featuring the expressed item (SEAL), a likely agent (SHARK), and a likely theme (FISH). Targets referred to the correct identity of the pronoun. Competitors referred to the incorrect identity of the pronoun.

<table>
<thead>
<tr>
<th>NP1 status</th>
<th>Construction</th>
<th>Sentence</th>
<th>Target</th>
<th>Competitor</th>
</tr>
</thead>
<tbody>
<tr>
<td>Expressed NP1</td>
<td>Active</td>
<td>The seal is quickly eating it</td>
<td>Likely theme</td>
<td>Likely agent</td>
</tr>
<tr>
<td>Expressed NP1</td>
<td>Passive</td>
<td>The seal is quickly eaten by it</td>
<td>Likely agent</td>
<td>Likely theme</td>
</tr>
<tr>
<td>Pronoun NP1</td>
<td>Active</td>
<td>It is quickly eating the seal</td>
<td>Likely agent</td>
<td>Likely theme</td>
</tr>
<tr>
<td>Pronoun NP1</td>
<td>Passive</td>
<td>It is quickly eaten by the seal</td>
<td>Likely theme</td>
<td>Likely agent</td>
</tr>
</tbody>
</table>

Importantly, prior research suggests that properties of NP1 can modulate comprehension of passive cues (Huang et al., 2013). In the Pronoun NP1 condition, NP1s are now likely agents in actives (e.g., “it” = SHARK) and likely themes in passives (e.g., “it” = FISH). Curiously, 5-year-olds from higher-SES backgrounds were as accurate with passives compared to actives in this context (58% vs. 59%). Huang and colleagues (2013) argued that unlike an expressed NP1, a pronoun NP1 generates a weak agent-first bias.¹ This delays role assignment until after the onset of the passive cue and NP2 (e.g., “…(eat)en by the seal”), allowing children to correctly infer that NP1 is a theme, without needing to revise a misinterpretation. Critically, to the extent that frequency differences between passives and actives do not impact interpretation when

¹ Recent work suggests that this effect is driven by the discourse status of NP1 and not general properties of the expressions, e.g., frequency, semantic sparseness (Huang, Abadie, Arnold, & Hollister, 2016). When children interpret pronoun NP1s as referring to new entities, they become more likely to recruit an agent-first bias and generate passive errors. Similarly, when expressed NP1s are contrasted with novel entities (e.g., “the blicker”), children consider them to be given entities and are less likely to apply an agent-first bias and produce passive errors.
revision is unnecessary, then a processing account predicts minimal SES-related differences in the Pronoun NP1 condition. In contrast, if challenges with passives reflect a failure to acquire syntactic structures in the first place, then a frequency account predicts that SES-related effects will emerge in this condition much like they do in the Expressed NP1 condition.

In addition to the language-processing task, we also assessed receptive vocabulary size. This measure served two primary purposes. First, since we did not directly quantify caregiver input, we needed to verify that SES-related differences in language abilities were present in our current sample. Input-related effects are well-established in vocabulary development (Arriaga et al., 1998; Hart & Risley, 1995; Hoff, 2003; Weisleder & Fernald, 2013), thus we used this as a proxy for variation in children’s language experience (for similar approaches, see also Fernald, Zangl, Portillo, & Marchman, 2008; Borovsky, Elman, & Fernald, 2012; Mani & Huettig, 2012). Second, we wanted to examine whether syntactic processing can shed light on relationships between SES background and vocabulary size. Admittedly, our cross-sectional design is not ideal for isolating causality. It could be that syntactic processing facilitates vocabulary growth, or input-related factors promote both vocabulary growth and syntactic processing. Nevertheless, in the area of word recognition, efficient processing of spoken utterances is positively linked to vocabulary size in both concurrent (Fernald et al., 2008; Borovsky et al., 2012; Mani & Huettig, 2012) and longitudinal measures (Fernald, Perfors, & Marchman, 2006; Marchman & Fernald, 2008; Weisleder & Fernald, 2013). This suggests that cross-sectional effects can provide some hints to the range of relationships that may exist throughout development.

2. Methods

2.1 Participants

One hundred and thirty-one English-speaking children participated in this study. We excluded data from two participants due to absence during the second testing session (n = 1) and
experimenter error during data collection (n = 1). This resulted in a final sample of 129 children (61 females, 68 males) with a mean age of 4;11 (SD = 0;9, range = 3;6 to 7;2). Children were recruited from local Head Start centers and private schools within the same geographical region. School status served as the initial basis for categorizing children’s SES background (65 lower SES, 64 higher SES). For 85% of the sample, detailed measures were also obtained through a questionnaire of parental education and annual family income, which were transformed from categorical variables into years of education and income in US dollars. When two parents had different levels of education or income, the higher of the two was selected. Among the sample, parents averaged 14.9 years of education (SD = 2.6 years; Range = 8 to 18 years) and averaged an income of $51,882 (SD = $35,378; Range = less than $15,000 to greater than $90,000).

Figure 1 illustrates the distribution of children’s ages across categories of family income.

![Figure 1](image_url)

**Figure 1.** A histogram of the average annual family income (in thousands of dollars) and age (in years) from 106 participants in the current study. This does not include information for 20 participants for whom a self-report survey was not returned.

Since prior research indicated that children had substantial difficulties when the NPI status of sentences varied across trials (Huang et al., 2013), we manipulated this factor between subjects. Sixty-three children were randomly assigned to the Expressed NPI condition and 66 children to the Pronoun NPI condition. Follow-up analyses confirmed that the two groups did
not differ in age, gender, school status, parental education, family income, or vocabulary size (all $p's > .30$). This ensured that differences in performance across NP1 status would not be caused by baseline differences in the general demographics of these two groups.

2.2 Procedure

For each participant, measures of language performance were obtained over two testing sessions. The language-processing task was administered during the first session, which lasted about 20-30 minutes. Children sat in front of an inclined podium divided into four quadrants, each containing a shelf where an object could be placed. On each trial, the experimenter labeled the objects in each set individually as they were placed on the shelf in a pre-specified order. This was followed by a pre-recorded sentence describing an event. Children were then encouraged to pick up the objects and use them to act-out what was said. Once the participant did this, the trial ended and the objects were returned to their pre-specified locations on the shelf. This was followed by a second pre-recorded sentence describing another event involving the same objects. Once children performed this action, the objects were removed from the display, and the next trial began with a new set of objects.

During the second session, receptive vocabulary was measured using the Peabody Picture Vocabulary Test-IV (PPVT; Dunn & Dunn, 2007). This lasted about 20 minutes. Testing items were divided into sets with 12 items each. For each item, children saw an array of four pictures and were asked to point to the one requested by the experimenter. They began with the set corresponding to their chronological age and stopped when they answered eight items in a set incorrectly. Since all our analyses statistically controlled for effects of age (see 3. Results section below), we measured vocabulary size using raw PPVT scores.

2.3. Materials
Critical trial types for the language-processing task represented the cells of a 2 x 2 design. The first factor, construction type, contrasts active versus passive sentences and was varied within subjects. The second factor, NP1 status, contrasts an expressed noun (e.g., “the seal”) versus pronoun (“it”) in the subject position and was varied between subjects. Visual displays featured three-object sets that paired the expressed item (e.g., SEAL) with a likely agent (e.g., something that can plausibly act on the expressed item, like a SHARK) and a likely theme (e.g., something that the expressed item can plausibly act on, like a FISH). The size of the items was controlled to ensure the plausibility of the relationship. Likely agents were always larger than expressed items, which in turn were larger than likely themes. Predicted relationships across these items were independently confirmed through separate norming data (see Huang et al., 2013 for more details). Across trials, each object type appeared in each location 33% of the time to ensure that the role of the object could not be predicted based on the display arrangement.

For each object set, we constructed critical sentences like those in Table 1. All sentences mentioned an expressed noun and a pronoun but varied in the order in which these occurred. Verb morphology distinguished between actives (i.e. present progressive) and passives (i.e., past participle). A be-auxiliary and adverb (e.g., “is quickly”) were embedded between NP1 and the verb to create a period of ambiguity when role assignments could not be informed by the event. Sentences were pre-recorded by a female actor who spoke in slow and consistent manner. Four versions of each item were used to create four presentation lists, such that each list contained six

---

2 It is well-documented that get-passives are more frequent than be-passives among adults from lower-SES backgrounds (Weiner & Labov, 1983; Sneller & Fisher, 2015). While future work will investigate these differences, additional factors led us to first focus on be-passives. First, get-passives are more semantically restrictive (McEnery & Xiao, 2005), thus they were not felicitous over a wide range of stimuli items. Second, get-passives are less likely to occur with adverbial modifiers (Carter & McCarthy, 1999), which were included in the current study. Finally, SES-related differences in the frequency of be-passives are consistent with the linking assumptions behind predictions by the frequency and processing accounts (pg. 7-8).
items in each condition and each item appeared just once in every list (see Appendix A for a full list of items). Twelve critical trials were pseudo-randomly presented with 32 additional filler trials that diverted attention away from the manipulated variables. To avoid systematically biasing the interpretation of NP1 as the agent or theme, filler sentences recruited symmetric predicates (e.g., “dance,” “fight”), experience and stimulus verbs (e.g., “like,” “scare”), and agent/theme intransitives (e.g., “sing,” “break”) (see Huang et al., 2013 for more details).

2.4 Coding

Approximately 1.3% of trials was excluded from subsequent analyses because of experimenter error. Data from all other trials were coded in the following manner.

**Fixations.** Trained research assistants coded eye movements using the frame-by-frame annotation software, Vcode (Hagedorn, Hailpern, & Karahalios, 2008). Trials began at the onset of the instruction and ended with the onset of the corresponding action. While blind to object locations and trial conditions, changes in gaze direction were coded as looks towards one of the quadrants, at the center, or missing due to looks away from the display or blinking. Missing frames accounted for 10.9% of coded frames. Remaining looks were then recoded based on their relation to the trial condition (see Table 1). “Target fixations” were defined as looks to pronoun referents that were consistent with correct role assignment. For Expressed NP1 / Passive and Pronoun NP1 / Active trials, this was the likely agent. For Expressed NP1 / Active and Pronoun NP1 / Passive trials, this was the likely theme. “Competitor fixations” were defined as looks to pronoun referents that were consistent with incorrect role assignment. For Expressed NP1 / Passive and Pronoun NP1 / Active trials, this was the likely theme. For Expressed NP1 / Active and Pronoun NP1 / Passive trials, this was the likely agent. Twenty-five percent of trials were checked by a second coder who confirmed the direction of fixation for 92.1% of coded frames. Disagreements between the two coders were resolved by a third coder.
Actions. Research assistants also coded videotapes of actions and categorized responses based on the trial condition. “Correct actions” were defined as those that depicted correct role assignments between the expressed item and Target. For Expressed NP1 / Passive and Pronoun NP1 / Active conditions, this referred to actions where likely agents did something to expressed items (e.g., making the SHARK eat the SEAL). For Expressed NP1 / Active and Pronoun NP1 / Passive conditions, this referred to actions where expressed items did something to likely themes (e.g., making the SEAL eat the FISH). “Reverse actions” were defined as those that depicted incorrect role assignments between the expressed item and Competitor. For Expressed NP1 / Passive and Pronoun NP1 / Active conditions, this referred to actions where expressed items did something to likely themes. For Expressed NP1 / Active and Pronoun NP1 / Passive conditions, this referred to actions where likely agents did something to expressed items. “Ambiguous actions” were defined as incorrect actions where the expressed item was selected with no additional object, all three objects were selected, or no object was selected at all.

3. Results

Since our procedures did not include direct measures of caregiver input, it was important to first establish whether SES-related differences in language abilities was present in our current sample. Consistent with this assumption, we found that vocabulary size correlated with family income ($r(101) = .49, p < .001$), parental education ($r(101) = .45, p < .001$), and school status ($r(101) = .46, p < .001$), while controlling for age. Given these robust relationships with a global measure of language ability, we isolated finer-grained effects in the following way. First, we examined fixations during spoken sentences to assess on-line sensitivity to the syntactic cues that distinguish actives and passives. Second, we examined actions following the sentence to assess likelihood of revising initial syntactic interpretation. Third, we returned back to the SES-related
effects of vocabulary development illustrated above and examined the extent to which child-to-child variation in vocabulary size is associated with syntactic-revision abilities.

Throughout our analyses, family income was used as a continuous measure of SES. This variable was highly correlated with parental education ($r(103) = .79, p < .001$) and school status ($r(106) = .91, p < .001$). Overall effects remained the same regardless of which measure was used, but they were most robust with family income as the predictor. Also, for ease of illustrating patterns (Figures 4 and 5), we categorized children into two groups based on a median split of family income (less/greater than $30,000) when these data were available and school status when they were not. Unless otherwise noted, dependent measures were analyzed separately across NP1 status, using logistic mixed-effects models. Subjects and items were included as random-effects variables (intercepts only). Construction type and age were included as fixed-effects variables. Gender was omitted since it never improved model fit (all $p$’s > .30). Analyses were conducted using the lme4 software package in R (Bates, Maechler, Bolker, & Walker, 2013).

3.1 Fixation analyses

To assess on-line sensitivity, we time-locked fixations to when pronoun referents could be identified via linguistic cues in the speech stream. This corresponded to the onset of the verb morphology in the Expressed NP1 condition (e.g., “eating” vs. “eaten”) and the onset of NP2 in the Pronoun NP1 condition (e.g., “seal”). Regions of analyses continued until 500 ms after sentence offset, generating an average window of 1023 ms in the Expressed NP1 condition and 1122 ms in the Pronoun NP1 condition. Also, preliminary analyses revealed that latency to look at the expressed item (e.g., SEAL) following the onset of the expressed noun (e.g., “seal”) was approximately 330 ms among a group of 40 adults but delayed until 530 ms in the current sample of children. We reasoned that any delays in restricting reference for expressed nouns would have cascading effects on role assignment and postpone looks to likely agents and likely themes (see
Huang et al., 2013 for a similar approach). To account for this developmental difference, regions of analyses were shifted by 400 ms after the onset of the input in the speech stream (i.e., time between adults and children plus the standard 200 ms estimate of how long it takes adults to generate a saccadic eye-movement – see Matin, Shao, & Boff, 1993).

Figure 2. Among children from higher-SES backgrounds, fixations to the expressed item (SEAL), likely agent (SHARK), and likely theme (FISH) after verb morphology in (A) Expressed NP1 / Active and (B) Expressed NP1 / Passive trials, and after NP2 in (C) Pronoun NP1 / Active and (D) Pronoun NP1 / Passive trials.

Figures 2 and 3 illustrate that children often looked to the expressed item, accounting for 44.4% of sampled fixations. To compare across conditions, we converted average, continuous fixations during the critical region of each trial into a binary variable (Jaeger, 2008). This binary variable accounts for the saccadic nature of eye-movements and best captured the underlying
distribution of our data. While eye gaze was sampled every 33 ms, children typically made only one or two saccades in a single second. Consequently, any measure of fixation proportion within that window is essentially binary for each trial. If average fixations during this region were greater than .50, then values were coded as 1. If they were less than .50, then values were coded as 0. Approximately 4.5% of trials were excluded because of no looks to the expressed item or no preference in either direction (i.e., average fixations of exactly .50). Analyses revealed no effects of age, income, or construction type on fixations to the expressed item in Expressed NP1 and Pronoun NP1 conditions (all z’s < 1.00, all p’s > .15).

Figure 3. Among children from lower-SES backgrounds, fixations to the expressed item (SEAL), likely agent (SHARK), and likely theme (FISH) after verb morphology in (A) Expressed NP1 / Active and (B) Expressed NP1 / Passive trials, and after NP2 in (C) Pronoun NP1 / Active and (D) Pronoun NP1 / Passive trials.
Importantly, Figures 2 and 3 also illustrate that Target and Competitor fixations varied by condition. After the onset of verb morphology in the Expressed NP1 condition (panels A and B), there was a preference for likely theme in active trials and likely agent in passive trials. This pattern appropriately switched after the onset of NP2 in the Pronoun NP1 condition (panels C and D). To compare linguistic and SES effects, we calculated preference scores for each trial. For passives, we subtracted Target minus Competitor looks, such that more positive values indicated greater sensitivity to passive cues. For actives, we subtracted Competitor minus Target looks, such that more negative values indicated greater sensitivity to active cues. We again converted average preference scores into a binary variable. If scores were greater than 0, values were coded as 1. If scores were less than 0, values were coded as 0. Approximately 15.7% of trials were excluded because of no Target or Competitor looks or equal looks to both objects. Follow-up analyses confirmed that the number of excluded trials did not differ by age, income, or construction type (all z’s < 1.50, all p’s > .10).

As expected, Figure 4a illustrates greater preference scores in passive trials compared to active trials in the Expressed NP1 condition. Importantly, SES-related differences also emerged. Target fixations following passives increased for children from higher-SES backgrounds (leading to scores above zero) but remained unchanged for those from lower-SES backgrounds (leading to scores around zero). This led to a critical interaction between family income and construction type (z = 2.31, p < .05), such that differences across actives and passives were greater in children from higher-SES backgrounds than those from lower-SES backgrounds. Planned comparisons revealed that family income was associated with more positive preference score for passives (z = 1.37, p > .15) and more negative scores for actives (z = 1.52, p > .10). However, these effects did not approach significance. Overall, there was an additional marginal effect of family income (z = 1.65, p < .10), but no main effects of age and construction type (all p’s > .20).
Figure 4. Fixation preference scores (A) after verb morphology in the Expressed NP1 condition and (B) after NP2 in the Pronoun NP1 condition. Correct fixations to the Target are indicated by positive scores in passive trials and negative scores in active trials.

Similarly, Figure 4b illustrates that preference scores in the Pronoun NP1 condition also increased in passive trials and decreased in active trials. These differences were again greater among children from higher-SES backgrounds than those from lower-SES backgrounds. This led
to a critical interaction between family income and construction type ($z = 2.63, p < .01$). Planned comparisons revealed that family income was associated with significantly more positive preference score for passives ($z = 1.98, p < .05$) and marginally more negative scores for actives ($z = 1.84, p < .10$). Overall, there was an additional marginal effect of family income ($z = 1.76, p < .10$), but no main effects of age and construction type (all $p$’s $> .20$). In sum, fixation patterns suggest SES-related effects on real-time syntactic processing. Irrespective of the need for syntactic revision, children from lower-SES backgrounds, on average, are less sensitive to the linguistic cues distinguishing actives and passives, compared to their higher-SES counterparts.

### 3.2 Action analyses

To examine how final interpretation varied with condition and SES background, we assessed the likelihood of correct responses in children’s actions. Figure 5 illustrates that overall accuracy was surprisingly low, even in the active trials (approximately 60%). Task complexity likely contributed to this effect. Each sentence included two NPs, an intervening adverb, and a pronoun whose identity had to be inferred based on the linguistic context. Importantly, active stimuli matched for these constraints. Average accuracy of actives was similar across Expressed NP1 (60%) and Pronoun NP1 conditions (55%), suggesting that task demands were matched across these contexts. Thus, they provide an appropriate baseline for the challenges associated with passives. Figure 5 also illustrates that errors with passives often involved role reversals, i.e., assigning *incorrect* roles to expressed nouns and selecting plausible pronoun referent on this basis. This suggests that failure to revise an agent-first contributed to errors in interpretation.

Figure 5a illustrates that all children were less accurate with passives compared to actives in the Expressed NP1 condition. However, those from lower-SES backgrounds found passives to be more challenging, on average, relative to their higher-SES peers. This led to an interaction between construction type and family income ($z = 2.22, p < .05$). Planned comparisons revealed
that family income was significantly associated with increased accuracy with passives ($z = 2.07$, $p < .05$), but marginally so with actives ($z = 1.70$, $p < .10$). Analyses also revealed main effects of age ($z = 2.06$, $p < .05$) and construction type ($z = 5.59$, $p < .001$), but not family income ($p > .30$). Importantly, Figure 5b illustrates that a different pattern emerged in the Pronoun NP1 condition. Similarities in accuracy across trials and SES background led to no effects of family income, construction type, or interaction between the two (all $p's > .20$). There was an overall main effect of age ($z = 4.06$, $p < .001$). In sum, action patterns suggest that children’s difficulties
with passives do not reflect a failure to acquire low frequency constructions. If this were the case, then the effects of construction type and SES background found in the Expressed NP1 condition should have also emerged in the Pronoun NP1 condition. Thus, their presence in the former but not the latter suggests that language experience impacts the retrieval of acquired representations, particularly in contexts that require syntactic revision.

Additional support comes from correlations between on-line sensitivity to linguistic cues and accuracy with passives. Recall that SES-related effects of on-line sensitivity were found in Expressed NP1 and Pronoun NP1 conditions. To quantify on-line sensitivity for each child, we calculated the average difference in preference scores for passive minus active trials. This value is greater for children who looked at correct Targets in passive (more positive scores) and active trials (more negative scores). Simple correlations revealed that on-line sensitivity was associated with passive accuracy in both Expressed NP1 \((r(63) = .33, p < .01)\) and Pronoun NP1 conditions \((r(66) = .29, p < .05)\). Partial correlations revealed that on-line sensitivity continued to predict passive accuracy in the Expressed NP1 condition, even when age, family income, and active interpretation were accounted for \((r(59) = .33, p < .01)\). However, the same was not true in the Pronoun NP1 condition \((r(61) = .13, p > .20)\). This suggests that SES-related differences in the comprehension of passives originate from factors associated with real-time processing. Variation in on-line sensitivity has minimal impact on comprehension accuracy when there is only a single interpretation (Pronoun NP1 condition). However, efficient access to syntactic representations may be critical when revising prior misinterpretation (Expressed NP1 condition).

### 3.3 Vocabulary analyses

Given the robust correlations between SES background and vocabulary measures in the current sample, we examined the extent to which finer-grained measures of syntactic processing can shed light on possible relationships between these factors. We reasoned that if SES-related
differences in language experience broadly impact language development, then vocabulary size may correlate more strongly with performance with passives compared to actives, irrespective of NP1 status. If, however, language experience impacts vocabulary development through specific relationships with syntactic processing, then we might expect correlations with vocabulary size to be more robust with passives that require syntactic revision compared to those that do not.

Table 2. In the Expressed NP1 condition, descriptive statistics and intercorrelations between tasks and demographics.

<table>
<thead>
<tr>
<th></th>
<th>FAMILY INCOME</th>
<th>ACTION (ACTIVE)</th>
<th>ACTION (PASSIVE)</th>
<th>VOCAB SIZE</th>
</tr>
</thead>
<tbody>
<tr>
<td>ACTIONS (ACTIVE)</td>
<td></td>
<td>.19</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>M = 0.60 SD = 0.26</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Range = 0 to 1.00</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ACTIONS (PASSIVE)</td>
<td></td>
<td>.38**</td>
<td>.36**</td>
<td>-</td>
</tr>
<tr>
<td>M = 0.35 SD = 0.31</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Range = 0 to 1.00</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>VOCABULARY SIZE</td>
<td></td>
<td>.45**</td>
<td>.35**</td>
<td>.45**</td>
</tr>
<tr>
<td>M = 87 SD = 25</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Range = 34 to 139</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AGE (MONTHS)</td>
<td></td>
<td>.27</td>
<td>.25*</td>
<td>.30*</td>
</tr>
<tr>
<td>M = 58 SD = 10</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Range = 43 to 86</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Note. Significant correlations marked with * for $p < .05$ and ** for $p < .01$

Table 3. In the Pronoun NP1 condition, descriptive statistics and intercorrelations between tasks and demographics.

<table>
<thead>
<tr>
<th></th>
<th>FAMILY INCOME</th>
<th>ACTION (ACTIVE)</th>
<th>ACTION (PASSIVE)</th>
<th>VOCAB SIZE</th>
</tr>
</thead>
<tbody>
<tr>
<td>ACTION (ACTIVE)</td>
<td></td>
<td>.15</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>M = 0.56 SD = 0.29</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Range = 0 to 1.00</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ACTION (PASSIVE)</td>
<td></td>
<td>.17</td>
<td>.61**</td>
<td>-</td>
</tr>
<tr>
<td>M = 0.53 SD = 0.28</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Range = 0 to 1.00</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>VOCABULARY SIZE</td>
<td></td>
<td>.62**</td>
<td>.43**</td>
<td>.32*</td>
</tr>
<tr>
<td>M = 86 SD = 29</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Range = 2 to 141</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AGE (MONTHS)</td>
<td></td>
<td>.22</td>
<td>.53**</td>
<td>.46**</td>
</tr>
<tr>
<td>M = 59 SD = 9</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Range = 42 to 84</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Note. Significant correlations marked with * for $p < .05$ and ** for $p < .01$
Tables 2 and 3 illustrate that vocabulary size was correlated with the accuracy of active and passive interpretation, age, and family income across levels of NP1 status (all $r’s > .30$, all $p’s < .01$). To test the robustness of these relationships, we conducted partial correlations that controlled for effects of age and performance with actives. Family income continued to correlate with vocabulary size in both Expressed NP1 ($r(48) = .36, p < .01$) and Pronoun NP1 conditions ($r(49) = .61, p < .001$). This demonstrates that SES-related effects on vocabulary development are broad and persist even when basic measures of cognitive development and comprehension are taken into account. In contrast, a different pattern emerged with passives. In the Expressed NP1 condition, passive interpretation continued to correlate with vocabulary size, even when effect of age and active interpretation were partialed out ($r(57) = .31, p < .05$). However, no relationship was found in the Pronoun NP1 condition, when other factors were accounted for ($r(58) = -.02, p > .80$). These results suggest that unlike measures of family income, performance with passives in the Expressed NP1 condition may isolate mechanisms that facilitate vocabulary development through efficient syntactic revision.

4. Discussion

This study investigated the causes of SES-related differences in syntactic development by examining the comprehension of actives and passives in children from diverse SES backgrounds. In measures of final interpretation, children across all backgrounds used linguistic cues to assign correct roles in actives and passives that did not require syntactic revision. However, they were less accurate when passives did require revision, and this was particularly true for children from lower-SES backgrounds, on average. Measures of real-time processing lend additional insight into these patterns. SES-related effects were found in on-line sensitivity to linguistic cues across contexts, but only predicted final interpretation when syntactic revision was needed. Altogether, these results suggest that SES-related differences in language experience facilitate retrieval of
syntactic representations during comprehension, enabling accurate interpretation in the face of parsing challenges. Yet, when these challenges are removed, effects of language experience also diminish, revealing strikingly similar comprehension abilities across SES background.

In the remainder of this discussion, we will focus on four additional issues related to the current findings. First, we will consider the extent to which differences in accuracy across NP1 status reflect statistical properties that exist in the input. Second, we will evaluate whether SES-related effects with passives can be explained by cognitive and/or linguistic factors that may vary in children from diverse backgrounds. Third, we will present some very preliminary evidence of what kinds of language experience may be relevant for explaining SES-related differences in interpreting passives. Finally, we will discuss the broader implications of the current findings for theories of language development and language processing.

4.1 Why are Expressed NP1 passives so hard to comprehend?

We have argued that challenges with passives in the Expressed NP1 condition reflect difficulties with revising initial role assignments. However, it is possible that the patterns we observed were instead driven by features of our task that were syntactically or pragmatically infelicitous. Here we consider two versions of this hypothesis. One possibility is that children’s challenges with passives were driven by their interpretation of the adverb rather than NP1. In particular, adverbs that encoded manner of motion may be more felicitous with actives, which typically convey on-going events (e.g., “quickly eating it”), but they may be infrequent with passives, which typically convey completed events (e.g., “quickly eaten by it”). Moreover, manner adverbs have been shown to increase children’s attention to event actions (Syrrett, Arunachalam & Waxman, 2014), thereby potentially increasing their agent-first bias.

We see three reasons to reject this account of our data. First, corpus analyses suggest that adverbs, including those that encode manner of motion, are not uncommon with be-passives. In
an analysis of spoken utterances from the British National Corpus (Aston & Burnard, 1998), McEnery and Xiao (2005) found that adverbs occurred in 19.5% of 5,001 instances of be-passives. Second, manner adverbs in the current study occurred in Expressed and Pronoun NP1 conditions. If they are infelicitous in passives, this should have negatively impacted accuracy in both cases. Finally, data from Mandarin passives suggest the agent-first bias is unaffected by the presence of adverbs (Huang et al., 2013). Unlike in English, the BEI marker disambiguates both arguments before the onset of the adverb (e.g., seal BEI it quickly eat → “The seal is quickly eaten by it”). Nevertheless, much like in the current study, Mandarin-speaking children were less accurate as using passive cues to identify pronoun NP2, when NP1 was an expressed noun. This suggests that the agent-first bias reflect properties of NP1 and not the adverb.

Table 4. Frequency of utterances categorized by NP1 status among 100 full be-actives (i.e., NP1 be VP-ing NP2) and be-passives (i.e., NP1 be VP-ed by NP2) in adult-direct utterances from the British National Corpus (Aston & Burnard, 1998).

<table>
<thead>
<tr>
<th>CONSTRUCTION</th>
<th>NP1 STATUS</th>
<th>EXPRESSED NP1</th>
<th>PRONOUN NP1</th>
</tr>
</thead>
<tbody>
<tr>
<td>ACTIVE</td>
<td>Example</td>
<td>One pigeon is eating the potato</td>
<td>He is kicking the ball</td>
</tr>
<tr>
<td></td>
<td>%</td>
<td>28</td>
<td>72</td>
</tr>
<tr>
<td>PASSIVE</td>
<td>Example</td>
<td>The woman was grabbed by a man</td>
<td>It is published by Oxford</td>
</tr>
<tr>
<td></td>
<td>%</td>
<td>57</td>
<td>43</td>
</tr>
</tbody>
</table>

A second possibility is that challenges with Expressed NP1 passives reflect frequency differences linked to the pragmatics of passives. Passives are used to highlight themes relative to agents (Johnson-Laird, 1968; Williams, 1977), and pronouns often refer to prominent discourse entities (Arnold, 2010; Arnold & Griffin, 2007), which usually occur in topic (NP1) position (Givon, 1983). These factors may lead to a frequency advantage for passives, featuring pronoun NP1s compared to expressed NP1s. To verify this, we randomly selected 100 instances of actives (i.e., NP1 be VP-ing NP2) and passives (i.e., NP1 be VP-ed by NP2) in adult-directed speech,
from the British National Corpus (Aston & Burnard, 1998). Among sentences similar to the current study, Table 4 confirmed that actives overwhelmingly featured pronoun NP1s compared to expressed NP1s (72% vs. 28%). In contrast, passives revealed an opposite preference for expressed NP1s compared to pronoun NP1s (57% vs. 43%). Thus, if a frequency advantage exists for pronouns NP1s, it is more likely to occur with actives and not passives.

Interestingly, similar patterns are also found in child production. Brooks and Tomasello (1999) asked 2- and 3-year-olds to generalize novel verbs in active and passive sentences, and their utterances were coded based on NP1 status (note that unlike the analyses above, children sometimes omitted NP1s in their utterances, thus percentages will not add up to 100%). Since production was elicited through question prompts, children generally preferred pronoun NP1s compared to expressed NP1s. Critically, this preference was the same across active (58% vs. 10%) and passive utterances (59% vs. 14%). In a follow-up experiment, the authors used a discourse manipulation to boost overall production of expressed NP1s. While actives still favored pronoun NP1s over expressed NP1s (45% vs. 29%), passives now featured comparable proportions of both (33% vs. 36%). Altogether, these data suggest that a lack of experience with expressed NP1 passives alone cannot account for our current patterns.

4.2 Why do SES-related differences emerge with passives?

We have argued that SES-related differences in interpreting passives reflect effects of language experience during the real-time processing of syntactic cues. However, it is possible that the patterns we observed were driven by other cognitive and/or linguistic factors that may vary with SES backgrounds. One possibility is that SES-related differences exist in the abilities to overcome basic task demands in the current study. Recall that critical sentences required children to correctly assign grammatical roles to the expressed noun (e.g., is “the seal” an agent or theme?) and use this knowledge to select a semantically plausible referent for the pronoun
(e.g., is “it” a fish or shark?). Thus, it may be the case that our findings do not reveal children’s challenges with syntactic revision, but instead reflect SES-related differences in more basic processes, like pronoun interpretation or semantic knowledge of related objects.

Yet, other features of children’s performance are inconsistent with this account. We found no SES-related differences in the accuracy of interpreting actives and passives that do not require revision. Like passives that required revision, these trials also involved utterances of a similar length, pronoun interpretation, and real-world knowledge. Perhaps most striking, children correctly inferred the referent of a pronoun NP1 in actives (e.g., “It is quickly eating the seal” → “It” is a SHARK), despite failing to do so when the same object was the referent of passives that required revision (e.g., “The seal is quickly eaten by it” → “It” is a SHARK). Moreover, even when children incorrectly interpreted these latter trials, their actions revealed an understanding of the basic task. They depicted far more plausible relationships (e.g., making the SEAL eat the FISH – lower: 57%, higher: 65%) than implausible ones (e.g., making the SEAL eat the SHARK – lower: 12%, higher: 10%), with no differences across SES background (all p’s > .20). This preference for plausible errors is consistent with failures in syntactic revision and suggests that real-world knowledge is unlikely to account for SES-related differences with passives.

A second possibility is that SES-related differences with passives reflect variation in the cognitive-control abilities that support real-time comprehension. Psycholinguistic research has revealed notable parallels between individual differences in syntactic revision and cognitive control tasks (e.g., Stroop, n-back), among experienced language users like adults (January, Trueswell, & Thompson-Schill, 2009; Kan et al., 2013; Novick, Hussey, Teubner-Rhodes, Harbison, & Bunting, 2014) and impaired populations like Broca’s aphasics (Novick, Kan, Trueswell, & Thompson-Schill, 2010). It has been argued that late maturation of cognitive-control abilities causes developmental challenges with syntactic revision (Novick, Trueswell, &
Thompson-Schill, 2005; Mazuka, Jincho, & Onishi, 2009). Importantly, executive function skills vary with SES backgrounds (Hackman & Farah, 2009; Blair et al., 2011), raising the possibility that the current findings reflect effects of cognitive control and not language experience.

Admittedly, it is difficult to address this account directly since we did not collect separate measures of cognitive control. Moreover, it may not be useful to consider effects of domain-general cognition as being mutually exclusive from those of domain-specific linguistic input, since accounts of language processing typically involve both (January et al., 2009; Kan et al., 2013; Novick et al., 2014; Novick et al., 2010; Novick et al., 2005; Mazuka et al., 2009). Indeed, recent research on bilinguals has highlighted ways in which language experience can shape cognitive-control abilities during development (Bialystok, Craik, Klein, & Viswanathan, 2004; Costa, Hernandez, Costa-Faidella, & Sebastian-Galles, 2009). Nevertheless, within the current study, asymmetries between on-line sensitivity and off-line interpretation suggest that SES-related effects have their basis in language experience. Even when comprehension accuracy was equivalent in the Pronoun NP1 condition, fixations to correct referents remained delayed among children from lower-SES backgrounds. This suggests that language experience impacts real-time processing even when syntactic revision is unnecessary.

4.3 What is the relationship between language experience and language interpretation?

While the current study sheds light on the potential mechanisms underlying SES-related differences in syntactic development, it leaves open what language experiences generate these effects. To tackle this question, future work will pair the current language-processing task with direct assessments of caregiver input. Nevertheless, given the current findings, we conducted preliminary analyses to explore the extent to which comprehension of actives and passives may be associated with a proxy of language experience: parental reports of the number of children’s books in the home. It is well documented that storybooks are a key source of linguistic input.
(Montag, Jones, & Smith, 2015) and a strong predictor of language outcomes (Senechal, LeFevre, Hudson, & Lawson, 1996). Among our current sample, families owned on average 51 books (SD = 32 books; Range = 0 to 80+ books). Unsurprisingly, book quantity was strongly correlated with family income ($r(106) = .67, p < .001$) and school status ($r(112) = .71, p < .001$).

Notably, when book quantity was added as a predictor of comprehension accuracy, its effects emerged specifically in the Expressed NP1 condition, and varied with family income and construction type ($z = 2.17, p < .05$). To unpack this 3-way interaction, we divided our sample into two SES groups (see pg. 16 for details) and correlated book quantity with accuracy in each condition, while controlling for age. In the passive trials, book quantity was positively associated with accuracy for children from higher-SES backgrounds ($r(24) = .26, p > .15$), but this did not approach significance. No relationship was found for their lower-SES counterparts ($r(23) = -.03, p > .80$). In contrast, in the active trials, book quantity was positively associated with accuracy for children from lower-SES backgrounds ($r(23) = .46, p < .05$). Curiously, this relationship did not emerge for their higher-SES counterparts ($r(24) = .12, p > .50$). Thus, consistent with prior work, these patterns confirm that language experience (as measured by book quantity) is clearly related to language outcomes (as measured by comprehension accuracy).

However, they also suggest that these relationships are likely more complex than a simple frequency effect (e.g., more input always boosts the number of passives heard). Instead, input quantity may interact with differences in the structures produced by caregivers from varying SES backgrounds (Huttenlocher et al., 2007; Huttenlocher et al., 2010). Among higher-SES families, greater input may be associated with an increased frequency of passives, and may contribute to the relationship between input quantity and passive comprehension found above. In contrast, among lower-SES families, greater input may be associated with an increased frequency of actives, thus generating a relationship between input quantity and active comprehension. These
effects may also be contribute to recent evidence that children from higher-SES backgrounds prefer to learn from informants who used passive sentences, while their lower-SES counterparts prefer informants who used active sentences (Corriveau, Kurkul, & Arunachalam, 2016).

Clearly, these patterns will need to be confirmed with direct measures of caregiver input. However, to the extent that they prove reliable, it suggests that the statistical profiles of syntactic structures associated with increased input may vary across SES background. This has direct implications for current interventions that focus on increasing caregiver input among lower-SES families (e.g., Thirty Million Words Initiative, Providence Talks). Such approaches may improve comprehension of constructions that are already frequent in the input, but they may be less effective for addressing SES-related differences in complex/infrequent syntactic structures.

4.4 Implications for theories of language development and language processing

By examining variation during development through the lens of real-time processing, the current findings inform theories of language in multiple domains. With respect to acquisition, this study provides surprising evidence of early proficiency with a low-frequency construction. This pattern is consistent with prior work in Mandarin (Huang et al., 2013). As in English, BEI passives occur less frequently than BA actives. Moreover, cross-linguistic comparisons suggest that passives are ten times less frequent in Mandarin than in English (McEnery & Xiao, 2005). Yet, when syntactic revision was not required, children’s accuracy with passives did not differ from actives. This pattern parallels current SES-related effects on passives, which are generally rare in caregiver input (Maratsos et al., 1985; Gordon & Chafetz, 1990; Stromswold et al., 2002) and may be even more so among lower-SES families. Critically, we found no evidence of SES-related differences for passives that did not require revision. This suggests that children can acquire knowledge of syntactic constructions, even when the input is minimal.
However, the current findings also suggest that language experience can impact syntactic development in other ways. Despite their accuracy with passives that do not require revision, children, particularly those from lower-SES backgrounds, were far less successful when revision was necessary. Cross-linguistic comparisons again provide interesting parallels with SES-related effects (Huang et al., 2013). While linguistic cues for English passives are recruited for multiple purposes (e.g., “-ed/en” associated with past tense, by-phrase also marks locations), BEI is used exclusively to signal passives in Mandarin. Yet, despite high cue validity, Mandarin-speaking children faced similar challenges with passives that require revision. Overall, these patterns are consistent with well-documented developmental difficulties with syntactic revision (Choi & Trueswell, 2010; Trueswell et al., 1999; Hurewitz et al., 2000; Weighall, 2008; Omaki et al., 2014). However, while prior accounts of these effects have focused on the role of age-related cognitive maturation (Novick et al., 2005; Mazuka et al., 2009), our results suggest that language experience may also contribute to efficiently accessing linguistic cues for revision.

Finally, a developmental perspective provides insights into the basic architecture of the language-processing system (Trueswell & Gleitman, 2004; Snedeker & Huang, 2015). Our findings reveal that despite substantial differences in language experience and cognitive abilities, adults and children recruit strikingly similar comprehension strategies. Like adults, children do not wait until the end of utterances to generate syntactic interpretation. Instead, their agent-first bias demonstrates a tendency to incrementally assign grammatical roles based on the statistical properties of the language. Also, like adults, children face processing difficulties when their initial interpretation conflicts with cues that occur later in the utterance (Trueswell et al., 1999; Hurewitz et al., 2000; Weighall, 2008; Choi & Trueswell, 2010; Huang et al., 2013; Omaki et al., 2014). Critically, while processing challenges in adults are often captured by ephemeral delays in comprehension, those in children tend to linger in the form of incorrect interpretation.
5. Conclusion

This study examined the role of SES background on syntactic processing of spoken utterances in 3- to 7-year-olds. Fixation patterns revealed average SES-related differences in real-time sensitivity to linguistic cues distinguishing active and passive sentences. Decreased sensitivity generated specific challenges for interpreting passives that required syntactic revision (e.g., reinterpreting “the seal” as the theme after hearing “…eaten by it”), leading to SES-related differences in the final interpretation of utterances. Importantly, we found that all children were equally proficient with passives that did not require revision (e.g., interpreting “it” as the theme after hearing “…eaten by the seal”), suggesting that language experience did not alter the ability to acquire structures for passives in the first place. Altogether, these results suggest that SES-related differences are present in real-time syntactic processing, and variation along this dimension impacts the accuracy of utterance interpretation.
Acknowledgments

We are grateful to Lauren Abadie, Erin Hollister, and other members of the Language and Cognition Lab for their assistance in data collection and coding. We would also like to thank the families and administrators at the Meade Village Head Start Program (Anne Arundel County, MD), Azeeze Bates Head Start Program (Washington, DC), and Tubman Head Start Program (Columbia, MD) for taking part in this study. Portions of this work have been presented at the 39th annual Boston University Conference for Language Development. This work was supported by an NSF IGERT fellowship to KL and a UMCP ADVANCE grant to YTH and MR.
References


Appendix A. Sentences and objects on critical trials of the language processing task.

<table>
<thead>
<tr>
<th>Condition</th>
<th>Sentence (active/passive)</th>
<th>Expressed Item</th>
<th>Agent</th>
<th>Theme</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Expressed NP1 The boy is gently kicking/kicked by it</td>
<td>BOY</td>
<td>HORSE</td>
<td>BALL</td>
</tr>
<tr>
<td>1</td>
<td>Pronoun NP1 It is gently kicking/kicked by the boy</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>Expressed NP1 The towel is gently cleaning/cleaned by it</td>
<td>TOWEL</td>
<td>DRYER</td>
<td>PAN</td>
</tr>
<tr>
<td>2</td>
<td>Pronoun NP1 It is gently cleaning/cleaned by the towel</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>Expressed NP1 The firefighter is quickly rescuing/rescued by it</td>
<td>FIREFIGHTER</td>
<td>HELICOPTER</td>
<td>CAT</td>
</tr>
<tr>
<td>3</td>
<td>Pronoun NP1 It is quickly rescuing/rescued by the firefighter</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>Expressed NP1 The girl is happily feeding/fed by her</td>
<td>GIRL</td>
<td>MOTHER</td>
<td>BABY</td>
</tr>
<tr>
<td>4</td>
<td>Pronoun NP1 She is happily feeding/fed by the girl</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>Expressed NP1 The rabbit is slowly eating/eaten by it</td>
<td>RABBIT</td>
<td>WOLF</td>
<td>CARROTS</td>
</tr>
<tr>
<td>5</td>
<td>Pronoun NP1 It is slowly eating/eaten by the rabbit</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>Expressed NP1 The seal is quickly eating/eaten by it</td>
<td>SEAL</td>
<td>SHARK</td>
<td>FISH</td>
</tr>
<tr>
<td>6</td>
<td>Pronoun NP1 It is quickly eating/eaten by the seal</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>Expressed NP1 The dog is slowly chasing/chased by it</td>
<td>DOG</td>
<td>CAR</td>
<td>RABBIT</td>
</tr>
<tr>
<td>7</td>
<td>Pronoun NP1 It is slowly chasing/chased by the dog</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>Expressed NP1 The girl is tightly hugging/hugged by her</td>
<td>GIRL</td>
<td>MOTHER</td>
<td>BABY</td>
</tr>
<tr>
<td>8</td>
<td>Pronoun NP1 She is tightly hugging/hugged by the girl</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>9</td>
<td>Expressed NP1 The frog is quietly catching/caught by it</td>
<td>FROG</td>
<td>DOG</td>
<td>FLY</td>
</tr>
<tr>
<td>9</td>
<td>Pronoun NP1 It is quietly catching/caught by the frog</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>10</td>
<td>Expressed NP1 The boy is carefully lifting him up/lifted up by him</td>
<td>BOY</td>
<td>DAD</td>
<td>BABY</td>
</tr>
<tr>
<td>10</td>
<td>Pronoun NP1 He is carefully lifting the child up/lifted up by the child</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>11</td>
<td>Expressed NP1 The rock is loudly smashing/smashed by it</td>
<td>ROCK</td>
<td>HAMMER</td>
<td>EGG</td>
</tr>
<tr>
<td>11</td>
<td>Pronoun NP1 It is loudly smashing/smashed by the rock</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>12</td>
<td>Expressed NP1 The cat is easily scaring/scared by it</td>
<td>CAT</td>
<td>DOG</td>
<td>MOUSE</td>
</tr>
<tr>
<td>12</td>
<td>Pronoun NP1 It is easily scaring/scared by the cat</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>