Childhood SES affects anticipatory language comprehension in college-aged adults

Melissa Troyer¹ (mtroyer@ucsd.edu) & Arielle Borovsky² (borovsky@psy.fsu.edu)
¹Department of Cognitive Science, University of California, San Diego
9500 Gilman Dr., La Jolla, CA 92093
²Department of Psychology, Florida State University
1107 West Call St., Tallahassee, FL 32304

Abstract
Childhood socioeconomic status (SES) has a broad impact on cognitive development including nearly every aspect of language ability. In infancy, lower SES is associated with delays in real-time language processing skills, but it is not known whether or how this relationship carries into adulthood. We explore these questions by assessing the timecourse of anticipatory sentence interpretation in a visual-world eye-tracking task in college-aged adults from higher and lower SES backgrounds. While there were only subtle SES-related timing differences in anticipation of a sentence-final target noun, we found SES-related differences in looks to competitor items on the screen. Particularly, individuals from higher SES backgrounds showed relatively more looks to action-related competitors just prior to onset of the target noun. These findings suggest that early SES influences the dynamics of lexical activation during sentence processing even in adulthood and highlight the importance of early lexical input and experience for adult language skill.

Keywords: sentence comprehension, language processing, eye movements, socioeconomic status, individual differences, language acquisition

Introduction
Understanding spoken language involves rapid and flexible deployment of expectations about speech, shaped at least in part by real-time activation of event and semantic knowledge (e.g., Metusalem et. al 2012) which can vary tremendously according to individual experience. In children, real-time language processing is influenced by individual differences in linguistic experience tied to the quantity of parental input (Weisleder & Fernald, 2013). In addition, differences in the amount of child-directed speech may vary according to household socio-economic status (SES; Hart & Risley, 1995), which may drive differences in language processing even in infancy (Fernald, Marchman, & Weisleder, 2012). However, it is not yet known whether or how these SES-related influences on language processing persist into adulthood. We investigated this question by measuring real-time language processing performance as a function of childhood SES in college-aged adults.

Socioeconomic status is a construct defined by a number of factors related to income and environment, with higher SES generally indicating greater occupational, educational, and economic prestige (Krieger, Williams, & Moss, 1997). Importantly, children in lower-SES households are at greater risk than higher-SES peers for physical, emotional, or mental health issues and are less likely to succeed in school or at work (Duncan, Yeung, Brooks-Gunn, & Smith, 1998; Tracy et al., 2008). Socio-economic status is also correlated with measures of cognitive development, including selective attention, short- and long-term memory, and, in particular, executive function and language skill (Hackman & Farah, 2009; Neville et al., 2013).

Children from lower-SES backgrounds have different experiences from those of their higher-SES peers starting in the womb (Stiles, 2008). They are less likely to experience cognitively stimulating environments (e.g., access to books, toys, etc.; Bradley et al., 2001; Farah et al. 2008). In addition, the quantity and quality of speech to children varies tremendously as a function of SES (Hart & Risley, 1995; Hoff, 2003), such that children from lower-SES backgrounds hear fewer words and less complex speech than those from higher-SES backgrounds.

SES-related differences in language skills emerge in infancy (Halle et al., 2009), and these differences have important consequences for the development of speech processing skills. Fernald and colleagues (2012) assessed the vocabulary and lexical processing skills of children from diverse socioeconomic backgrounds between the ages of 18-24 months. Strikingly, they found that differences in both measures as a function of SES appeared even at 18 months. Consequently, long-range language learning trajectories associated with SES in childhood seem to be initially predicted by differences in the development of basic language processing skills from infancy.

Importantly, real-time language processing skills predict individual differences in both language and other cognitive abilities across childhood. In infancy, the speed and accuracy of lexical recognition are linked with current and future vocabulary skills and are associated with later cognitive outcomes at age 8 (Fernald, Perfors & Marchman, 2006; Marchman & Fernald, 2008). These associations scale up to more complex sentence processing tasks in toddlerhood (Mani & Huettig, 2012) and into childhood and adulthood (Borovsky, Elman, & Fernald, 2012; Borovsky & Creel, 2014). Reading and other receptive language skills are also associated with language processing abilities in school-age children (Nation, Marshall, & Altmann, 2003; McMurray, Munson & Tomblin, 2014). Therefore, an individual’s ability to recognize and interpret information from spoken language is likely both dependent on his or her early language environment and crucial for other skills that have far-
reaching consequences for academic and professional success.

In adults, evidence from behavioral measures like eyetracking (e.g., Kamide, Altmann, & Haywood, 2003) and electrophysiological measures like event-related potentials (e.g., Delong, Urbach, & Kutas, 2005) suggests that language processing continually involves pre-activation of likely upcoming linguistic material (see DeLong, Troyer, & Kutas (2014) for a recent review). Fluent listeners do not passively wait to receive information from the unfolding speech stream to begin comprehension. Instead, information from previous context allows listeners to generate predictions and pre-activate content prior to directly encountering it (e.g., Kamide et al., 2003). The visual world eye-tracking paradigm (VWP; Tanenhaus, Spivey-Knowlton, Eberhard, & Sedivy, 1995) is one of several methods that have been particularly productive in elucidating the predictive nature of language interpretation. The VWP has also differentiated between individual differences in adolescent and adult language processing in several groups, including healthy children and adults as well as adolescents with specific language impairment (SLI) (Borovsky et al., 2012; Borovsky, Burns, Elman, & Evans, 2013; Mani & Huettig, 2014; McMurray et al., 2014).

Here, we use the VWP to investigate differences in how adults from a range of SES backgrounds use context to anticipate upcoming linguistic content. Although prior work (reviewed above) indicates that SES has an impact on language-processing skills in infancy, little is known regarding whether this relationship continues into adulthood. One possibility is that young adults from lower SES backgrounds continue to show relative slowing in real-time language comprehension. Such a finding would indicate that the slowed language processing may contribute to the lifelong risks for other negative consequences associated with lower SES. It is also possible that SES-related differences in the pace of language processing disappear in adulthood, only to shift to difficulties in other, perhaps more subtle aspects of language processing. For example, adolescents with SLI do not differ from typically-developing peers in the speed or accuracy of anticipatory processing during sentence comprehension. Instead, they fail to activate less-likely sentence completions as the sentence unfolds, suggesting differences in real-time dynamics of lexical activation (Borovsky et al., 2013). Similarly, differences in the quantity of lifetime language experiences between higher and lower SES groups (Hart & Risley, 1995; Hoff, 2003) may shift the degree to which listeners activate uncertain or unexpected outcomes during linguistic processing.

We explore these issues by measuring the relationship between childhood SES and language processing skills in college-aged adults. Specifically, we compare performance on an eyetracking task that has previously highlighted differences in anticipatory speed and lexical activation during sentence processing (Borovsky et al., 2012, 2013).

**Methods**

**Participants**

145 college students from UCSD (N=50) and Florida State University (FSU) (N=95) participated for course credit. Participants were excluded from analysis if they reported a current or prior hearing or speech disorder or exposure to languages other than English during early childhood, leaving a total of 108 participants in the combined sample.

**Stimuli and design**

Linguistic materials were eight sentence quartets created by crossing two agents and two actions. All sentences had the same syntactic form: ‘The NOUN VERBs the NOUN.’ These sets were paired with four images related to the content of the sentences. Participants saw the four images concurrently as they heard each sentence unfold. Each image served a different purpose for each sentence in the quartet. For the sentence, ‘The pirate chases the ship,’ the Target item was the ship. Three distractors were the treasure (Agent-related; e.g., related to the agent noun pirate), the cat (Action-Related; e.g., a potential patient of the action verb chases), and the bones (Unrelated). Each image appeared once in each condition and was therefore able to serve as its own control (see Figure 1; all sentence materials are provided in Borovsky et al., 2012). Each participant heard 16 out of the total 32 sentences, and two sentences per quartet were heard by an individual participant so that each participant saw each image twice. Across participants, each object was presented an equal number of times in each condition and screen quadrant, and spoken sentences were constructed so that each word had the same length. (see Borovsky et al., 2012, for more information about the images and spoken sentence stimuli).

[Figure 1: Sample images and sentences]
Procedure

Participants were seated in front of a 17-inch LCD display. A standard 5-point calibration procedure was conducted and stimuli were presented using the EyeLink Experiment Builder software. Participants were instructed to view the images and to click the picture that “goes with the sentence.” As in previous work, we expected participants to click on the sentence-final noun, which was the Target. Participants completed one practice trial before beginning the experiment.

Before each trial, a bullseye fixation appeared which checked for excessive drift. The four-array image then appeared on the screen for 2000 ms before the sentence began. The images remained on the screen through the duration of the sentence until the participant clicked on a picture.

Eye movement recording

We used an EyeLink 2000 eye-tracker with remote arm configuration at 500 Hz for the data collected at UCSD and an Eyelink 1000+ remote eye-tracker with identical camera and data sampling configuration at FSU. We recorded eye movements for each trial beginning at the appearance of the image and ending with the mouse click. Data were binned offline into 50-ms intervals for further analysis.

Offline measurements of SES

We assessed maternal and paternal occupation and education using the Barratt Simplified Measure of Social Status (BSMSS; Barratt, 2006), an updated version of the four-factor Hollingshead (1975) SES measure. Because this score reflects parental measures of SES, we took this as a measure of an individual’s childhood SES.

Because SES scores for UCSD \((M = 47.97, SD = 12.54, \text{Range} = 25-66)\) and FSU \((M = 48.90, SD = 10.31, \text{Range} = 21-66)\) were similar, \(W = 1237.5, p = .88, r = -.014,\) we pooled the data from both groups \((N=108)\). Median splits determined Higher and Lower SES group membership. This led to the exclusion of three participants who scored exactly the median (resulting \(N=105\)).

Results

Accuracy

Correct responses were coded as trials where participants selected the picture that matched the sentence-final theme. Accuracy on the task was very high (99.3%), with 12 total errors out of 1717 recorded trials. Due to a computer error, an additional 9 responses were not recorded.

Eye-tracking timecourse characterization

Our first goal was to characterize the timecourse of fixations towards the target and distractor objects across the entire sentence for each SES group before carrying out statistical comparisons of looks towards the target in time windows of interest. We therefore calculated the mean proportion of time spent fixating to the Target, the Agent-Related distractor, the Action-Related distractor, and the Unrelated distractor in 50 ms bins from sentence onset to offset for Higher-SES and Lower-SES groups (Figure 2).

As in previous work, we observed a rapid anticipatory shift in looks to the Target following the onset of the verb. In addition, there was an increase in looks to the Action-Related distractor shortly after the onset of the verb as well as an increase in looks to the Agent-Related distractor following the onset of the subject noun. All of these tendencies replicate patterns found in previous work using these materials (Borovsky et al., 2012; 2013).

Our first analytic goal was to determine the time points when fixations towards the Target significantly exceeded those to the Agent-Related distractor for each of the Higher- and Lower-SES groups. For instance, given the sentence, ‘The pirate chases the ship,’ we were interested in identifying the moment when looks first diverged between the Target image (SHIP) and the Agent-Related distractor, (CHEST). As the sentence-initial agent (‘The pirate’) is spoken, there should be no reason to prefer one of these alternatives over the other, but as the sentential action is spoken (‘chases’), participants begin to make use of additional information available from the action to anticipate the Target completion (SHIP). To determine whether there were differences between SES groups in the
timing of looks to the target compared to competitors, we performed two separate analyses. First, we computed one-tailed t-tests over 50 ms bins on raw proportion data to determine the first point at which individuals reliably looked to the Target compared to the Agent-Related competitor. To meet this criterion, all subsequent bins over the sentential period needed to show the same pattern (more looks to Target than Agent-Related competitor). This analysis revealed that the Higher-SES group began to reliably look to the Target by 1150 ms post sentence onset (about 250 ms post action onset; $t(51) = 1.85, p < .05, d = .32$) whereas the Lower-SES group did not look reliably to the Target until 50 ms later, at 1200 ms post sentence onset (300 ms post action onset; $t(51) = 2.14, p < .05, d = .44$).

The second analysis was a more statistically rigorous direct comparison of SES differences in timing of looks towards the Target. We first transformed differences in raw proportion between looks to the Target and each of the distractors, respectively, to log-gaze probability ratios. Unlike raw proportion measures which are bounded between 0 and 1, this transformation defines the bias of looking to the Target relative to each of the other distractors and provides the benefit of allowing the dependent variable to range in value between positive and negative infinity (for further explanation and similar approaches, see Arai, van Gompel, & Scheepers, 2007). Next, we asked whether group differences in looks towards the Target relative to each type of distractor emerged by computing t-tests on log-gaze probability ratios between Higher- and Lower-SES groups in each 50 ms bin. There were no significant differences between groups at any time window when comparing total SES scores. There was a marginal effect of SES looks to the Target vs. Agent-Related distractor in the time window between 1100-1150 ms, $t(99.70) = 1.429, p = .08, d = .280$. This time period is just prior to the point at which the High-SES group began to converge on looking to the Target. Thus, both analyses suggest only modest timing differences between Higher- and Lower-SES groups.

Eye-tracking effects in periods of interest

Next, we asked whether the magnitude of fixations to each condition varied according to SES over two longer anticipatory time windows during (1) the action verb and (2) the following article. In these time periods, participants have encountered the information necessary to generate a prediction for the Target but have not yet heard the final noun. We calculated relative looking times in these time periods with log-gaze probability ratios of looks to the Target vs. Agent-Related distractor, Target vs. Agent-Related distractor, Target vs. Unrelated distractor. We compare differences across Higher- and Lower-SES groups using two-way t-tests in both time windows.

During the action region, there were no significant differences among SES groups in Target vs. Agent-Related or Target vs. Unrelated log-gaze probabilities, $p > .05$. However, there was a significant effect of SES group on the Target vs. Action-Related comparison: the Higher-SES group looked relatively more than the Lower-SES group toward the Action-Related distractor than toward the Target, $t(100.47) = -2.55, p < .05, d = -0.50$.

Results were similar during the article region. There were no significant effects of SES group on either Target vs. Agent-Related or Target vs. Unrelated comparisons, $p > .05$. Again, SES group had a significant effect on the Target vs. Action-Related comparison: the Higher-SES group looked more than the Lower-SES group toward the Action-Related distractor compared to the Target, $t(91.56) = -3.08, p < .01, d = -0.60$.

We also computed correlations between the composite SES score and each of the dependent measures defined previously: log-gaze probability ratios of looks to the Target vs. Agent-Related distractor, Target vs. Action-Related distractor, and Target vs. Unrelated distractor.

During the action time period, there was a significant negative relationship between SES and looks to the Target vs. Action-Related distractor, $r = -.24, p < .05$, indicating that individuals with higher SES scores were more likely to look toward the Action-Related distractor than individuals in the Lower-SES group. During the article time period, there was a significant negative relationship between SES and looks to the Target vs. Action-Related distractor, $r = -.26, p < .05$, indicating that individuals with higher SES scores were more likely to look toward the Action-Related item. There were no other significant correlations between SES and looking time comparisons in either time period. To illustrate the difference between looking times to the Action-Related vs. Unrelated distractor (as a baseline) for the Higher- and Lower-SES groups, we replot raw looking times to these two interest areas over the entire timecourse, including both groups on the same plot (Figure 3).

Discussion

This study investigated whether childhood SES background may manifest in language processing differences in adulthood using an eye-tracked sentence comprehension task. We initially outlined two potential hypotheses: (1) that SES may be linked with processing speed (as in infancy) or (2) that lexical dynamics/activation may vary according to SES (as in SLI). Our findings lend minimal support to the first hypothesis and strongly support the second.
We see relatively small differences in the timing of anticipatory looks to the target; however, the differences that do exist are in the expected direction. Participants from higher-SES backgrounds were only slightly faster than those from lower-SES backgrounds to look at the Target. This finding indicates that SES-related differences in the linguistic processing speed that exist in infancy extend into adulthood, but this effect is relatively small. Individuals from lower-SES backgrounds appear to “catch up” in terms of speed, at least in this relatively simple task.

Instead, our findings lend greater support for a lexical activation account. We found that individuals from higher-(vs. lower-) SES backgrounds showed relatively more robust looks to unexpected but potentially plausible endings that cohere with the local semantic content. Although this finding is somewhat unexpected, it is consistent with prior findings that adolescents with SLI are less likely than typically-developing peers to look toward action-related competitors in an identical task (Borovsky et al., 2013).

One possibility for the spike in looks to action-related competitors is that individuals may temporarily entertain locally coherent but globally unexpected linguistic information as a safeguard in case of encountering unexpected information in sentence comprehension. This explanation is consistent with connectionist models like TRACE (McClelland & Elman, 1986) that allow for temporary activation of items consistent with local context but inconsistent with prior information.

Language comprehension involves the dynamic activation of words and concepts in response to a constantly changing speech stream, and the shape of this activation is modulated by individuals’ knowledge and experience. Therefore, SES-related variability in the quantity and quality of early language experience may affect the breadth of lexical items likely to be activated at any given point during language comprehension. Because individuals from lower-SES backgrounds have likely been exposed to fewer words relative to higher SES peers, they may be less likely to pre-activate multiple lexical items for that context. This possibility has to do with the probabilistic distribution of lexical items in a given context and not with total vocabulary per se. For instance, following a sentence beginning, The boy is reading the..., an individual experiencing a wealth of linguistic input might hear words like book, magazine, novel, poem, story, and so on whereas an individual experiencing a relatively lower level of linguistic input might only encounter book in this context. It may therefore make sense for the former individual to “hedge their bets” and entertain less likely outcomes whereas an optimal strategy for the latter individual would be to stick with the most likely option (e.g., book).

Our speculation that adults from lower-SES backgrounds (pre-)activate fewer lexical items in language comprehension leads to the prediction that these adults will be impaired in situations where less-likely or novel information occurs. For instance, interpretation of so-called “garden-path” sentences or other ambiguous content in language may be more difficult for individuals who are less likely to maintain multiple lexical (or syntactic) representations at once. We hope to investigate these hypotheses in future research.

While our findings do suggest that childhood SES continues to exert an influence on adult language processing skills, we should note some important limitations regarding our sample. First, although we attempted to recruit a large and diverse sample at two geographically distinct public institutions across the United States, we should note that our participants are nevertheless attending selective college institutions and gained entry partially by achieving requisite language scores on major standardized college entrance examinations. It is likely that recruiting a community sample of adults of similar ages would increase the range of socio-economic background of our participants as well as variability in real-time language processing performance on this task. It is nevertheless notable that, despite this restriction in SES variability in our sample, we still find SES-related differences in language processing skills.

A secondary limitation is that we are using a relatively simple task that was designed to be easily understood even by preschool participants. It is possible that we will find more robust differences in timing with a more challenging language task that uses more advanced vocabulary or grammatical structures.

Despite these limitations, our findings highlight at least one way in which social inequities in childhood may impact adult language function. Early SES-related differences in the language experiences of children clearly have an impact in how adults navigate real-time language and develop expectations for speech. These changes in the dynamics of lexical activation may have important consequences for how early SES may affect older child and adult learners when they encounter novel or unexpected information, although further work is needed to understand precisely how this may occur. A greater implication of our findings is that early intervention in language skills is likely to have important positive implications even into adulthood for outcomes in the lives of children from lower-SES backgrounds.

Acknowledgments

We thank Tessa Arsenault and Virginia Prestwood for their assistance with collection and data processing. This work was supported by grants F32 DC010106 and R03 DC013638 to AB and NSF Graduate Research Fellowship and Kroner Fellowship to MT.

References


Hollingshead, A.B. (1975). Four factor index of social status. Unpublished manuscript, Department of Sociology, Yale University.


