



ELSEVIER

Contents lists available at ScienceDirect

Journal of Experimental Child Psychology

journal homepage: www.elsevier.com/locate/jecp



Brief Report

Developmental parsing and linguistic knowledge: Reexamining the role of cognitive control in the kindergarten path effect

Yi Ting Huang*, Erin Hollister

Department of Hearing and Speech Sciences, University of Maryland–College Park, College Park, MD 20742, USA

ARTICLE INFO

Article history:

Received 27 July 2018

Revised 5 April 2019

Available online 28 April 2019

Keywords:

SES differences

Sentence processing

Garden path recovery

Passives

Cognitive control

Linguistic knowledge

ABSTRACT

During spoken language comprehension, young children have difficulties in revising incorrect predictions about the structural properties of sentences. Recent research on individual differences suggests that these errors may reflect immature cognitive control. However, this evidence overlooks challenges with interpreting cross-task correlations and additional effects of linguistic knowledge on developmental parsing. To account for within-individual variation in task performance, this study compared sentence comprehension across two samples: one where socioeconomic status (SES) background was related to global language knowledge (Diagnostic Evaluation of Language Variation–Screening Test; $n = 60$) and another where it was related to cognitive control abilities (Stroop task; $n = 46$). Children (3- to 6-year-olds) heard sentences with an agent-first bias (e.g., “The blicket will be quickly ...”) that predicted late-arriving verb morphology (e.g., actives: “... eating the seal”) or conflicted with this cue (e.g., passives: “... eaten by the seal”). Consistent with prior work, final interpretations were less accurate when revision was needed for passives compared with actives. Critically, when SES was a proxy for global language variation, children from higher-SES backgrounds revised mispredictions more than their lower-SES peers on average. However, when SES tracked variation in cognitive control instead, SES effects on revision were absent. This suggests that variation in revising mispredictions during development may be related to linguistic knowledge rather than cognitive control. We discuss these results in light of known effects of linguistic knowledge on sentence comprehension and describe an

* Corresponding author.

E-mail address: ythuang1@umd.edu (Y.T. Huang).

information-theoretic account for why limited knowledge may lead children to favor predicted meanings over revised meanings during comprehension.

© 2019 Elsevier Inc. All rights reserved.

Introduction

During spoken language comprehension, children anticipate meanings based on early-arriving cues but often fail to revise predictions that turn out to be wrong. This is dubbed the kindergarten path effect. For example, in sentences like Sentence 1 below, knowledge of canonical word order leads 3- to 7-year-olds to interpret initial arguments as agents (i.e., first noun phrases [NP1s] are doers of actions). This leads to accurate interpretation of actives but hinders passives, which requires revision after verb morphology (Huang & Arnold, 2016; Huang, Leech, & Rowe, 2017; Huang, Zheng, Meng, & Snedeker, 2013). The kindergarten path effect is pervasive. Children often fail to revise mispredictions from verb biases (Kidd, Stewart, & Serratrice, 2011; Trueswell, Sekerina, Hill, & Logrip, 1999), case marking (Choi and Trueswell, 2010), and filler-gap dependencies (Omaki, Davidson White, Goro, Lidz, & Phillips, 2014). However, they readily interpret equivalent constructions without temporary ambiguity (Huang et al., 2013; Hurewitz, Brown-Schmidt, Thorpe, Gleitman, & Trueswell, 2000). This has promoted an assumption that the kindergarten path effect does not reflect inadequate linguistic knowledge.

1. *Active*: The blicket will be quickly eating the seal. [*blicket* is an agent]
2. *Passive*: The blicket will be quickly eaten by the seal. [*blicket* is a theme]

Instead, a prominent hypothesis has been that revision errors arise from immature cognitive control, which limits suppression of dominant but incorrect syntactic parses (Novick et al., 2005; Mazuka, Jincho, & Onishi, 2009). Recently, Woodard, Pozzan, and Trueswell (2016) tested this hypothesis by assessing 5- and 6-year-old children on a battery of language and cognitive tasks. In temporarily ambiguous sentences like “Put the frog on the napkin into the box,” children often construed first prepositional phrases (PP1s) as goals (i.e., “Put it on the napkin”) and maintained this misprediction after PP2s revealed that PP1s are in fact modifiers (i.e., “frog that’s on the napkin”). Importantly, individual variation in syntactic revision correlated with performance on a nonsyntactic conflict task. Children who made fewer revision errors also generated faster response times (RTs) and fewer errors when switching from congruent trials to incongruent trials of a Flanker task (see Qi, Fisher, & Brown-Schmidt, 2011, for analogous patterns from a Simon Says task). Based on these results, Woodard et al. (2016) suggested that “the ability to revise initial interpretive commitments is supported by domain-general executive function abilities, which are highly variable and not fully developed in children” (p. 187).

However, this conclusion hinges on the interpretability of correlational analyses. Because correlations track ranked performance across tasks (e.g., sentence vs. Stroop), they require measures with low within-individual variance (i.e., reliable) and high between-individual variance (i.e., people differ in performance). A classic example of this is standardized assessments. To distinguish impairment status in children, these tests adopt scales where each score equally informs how one child differs from another. In contrast, experimental tasks typically isolate shared processes; thus, they are crafted to yield homogeneous performance (e.g., everyone shows Stroop conflict). Yet, tasks with low between-individual variance are ill-suited for ranking individual performance (i.e., low intraclass correlation) because doing so requires variation to exist in the first place (Hedge, Powell, & Sumner, 2018). Moreover, experimental tasks, particularly those of the developmental variety, often exhibit high within-individual variability. For example, in a six-trial experiment, participant means of 0% and 100% have no performance variation ($SD = 0\%$), whereas those with 33%, 50%, and 67% have substantially more ($SD = 52\text{--}55\%$). This suggests considerable uncertainty about the underlying abilities in the latter group.

Yet, it is regularly assumed that variation in participant means directly corresponds to between-individual differences in abilities (e.g., a child with $M = 67\%$ is better than a child with $M = 50\%$, who in turn is better than a child with $M = 33\%$). This implies certainty that is unlicensed by these measures.

To isolate relationships between cognitive control and the kindergarten path effect, this study took a different approach. Rather than correlating individual performance, it examined distinctions that emerge across groups on average. It is well known that global language abilities (Hart & Risley, 1995; Hoff, 2003; Rowe, 2008) and cognitive control abilities (Noble, McCandliss, & Farah, 2007; Noble, Norman, & Farah, 2005) differ in children across socioeconomic status (SES) backgrounds. Moreover, when global language is assessed through measures like the Diagnostic Evaluation of Language Variation–Screening Test (DELV-S; Seymour, Roeper, & De Villiers, 2003), it predicts SES variation in syntactic revision (Leech, Rowe, & Huang, 2017). In sentences with a strong agent-first bias (e.g., definite NP1s: “The seal is eating/eaten by it”), children from lower-SES families on average are less likely to revise passives compared with their higher-SES peers (SES \times Construction, $p < .05$). However, when the agent-first bias is weakened and revision is unnecessary (e.g., pronoun NP1s: “It is eating/eaten by the seal”), children accurately interpret actives and passives with no SES effects ($p > .80$). This suggests that global language may track how children recover from mispredictions (see also Huang et al., 2017, for relations to vocabulary size). Critically, unlike cross-task correlations, this study isolated individual differences through mixed-effects models. By including subject as a random-effects variable, these analyses account for within-individual variance across trials.

However, prior work did not assess cognitive control abilities; thus, it remain unknown how this dimension contributes to the kindergarten path effect. To address this question, the current study took advantage of the fact that SES variation reflects *average* differences across populations. Thus, equivalent performance can be found within diverse samples. For example, in 38 children with slightly below-average language abilities, Noble, Wolmetz, Ochs, Farah, and McCandliss (2006) found no SES differences in vocabulary size, phonological awareness, and word reading measures. Likewise, the current study compared syntactic revision in Leech et al. (2017) with a new sample where SES was related to cognitive control (i.e., Stroop task) but *unrelated* to global language (i.e., DELV-S). This approach solved two problems that are common to correlational analyses. First, it is often unclear whether correlations reflect properties that are general to populations or specific to samples. By comparing groups with different profiles, we provided a baseline for interpreting co-occurring abilities (e.g., does global language still relate to revision when SES effects are minimal?). Second, by recruiting SES as a proxy for between-individual variability in global language or cognitive control, we adopted mixed-effects models to account for within-individual variability in comprehension.

Given this design, our predictions are as followed: If the kindergarten path effect reflects immature cognitive control, then SES differences in revision may emerge when SES effects on cognitive control are found. This would mirror global language patterns. If, however, prior correlations do not adequately account for within-individual variability in task performance, then SES differences in revision may be minimal when cognitive control varies but global language does not.

Method

Participants

A total of 50 children were recruited from schools in the Washington, D.C. metro area in the eastern United States. Data from 1 child was excluded due to absence, data from another child was excluded due to equipment failure, and data from two more children were excluded because English was not their primary language. This yielded a final sample of 46 children (mean age of 4;8 [years;months], $SD = 0;7$, range = 3;9–6;3). For 93% of the sample, demographics were obtained through a parental questionnaire. Parents averaged 15 years of education ($SD = 3$, range = 11–18) and family incomes averaged \$67,115 ($SD = \$36,497$, range = <\$15,000 to >\$90,000). Compared with Leech et al. (2017), family income was higher in the current sample ($p < .01$), but relationships between income and parental education were present across samples ($p < .001$) to a similar degree ($p > .90$). Children's age did not vary with family income or study ($ps > .15$).

Procedures and material

Global language and cognitive control measures were collected in the first session, and syntactic parsing was collected in the second session (~20–30 min each). Global language was assessed using the risk status subtest of the DELV-S. Children saw displays paired with prompts and were tested on their knowledge of “wh-” movement, auxiliaries, copulas, and pronouns. Cognitive control was assessed through a child-friendly Stroop task (Beveridge, Jarrold, & Pettit, 2002). Children were taught the names of four dogs, and then encountered 16 trials where a random dog appeared on the display. Their task was to say the dog’s name as quickly and accurately as possible. Congruent trials involved a green dog named *Green* and a brown dog named *Brown*. Incongruent trials involved a blue dog named *Red* and a red dog named *Blue*.

Syntactic parsing was assessed using a word learning task developed by Huang and Arnold (2016). Trials began with an animated event where a novel agent acted on a familiar object (e.g., a big scary monster chasing a seal) and a familiar object acted on a novel theme (e.g., a seal chasing a wimpy creature). This was followed by a sentence like “The blicket will be quickly *eating/eaten* by the seal.” Children then selected an object for the novel word. Previous work demonstrates that novel NP1s incur a strong agent-first bias, which improves accuracy for actives but hinders passives. Importantly, children interpret constructions equally well when the agent-first bias is weakened.¹ For each item, actives and passives were divided across two lists. Each list contained six items per construction, with each item appearing once in every list. Within lists, critical trials were randomized, with six filler trials featuring known words in active sentences.

Results

The data were analyzed with the lme4 software package in R (Bates, Maechler, Bolker, & Walker, 2015). Maximal mixed-effects models included random slopes and intercepts for participants and items. When these failed to converge, simpler models were adopted with random intercepts only (Barr, Levy, Scheepers, & Tily, 2013). Across analyses, family income was a continuous measure of SES, and age (in months) was added as a predictor, except for within-condition analyses of experimental tasks (i.e., Stroop, Sentence) where age and family income were collinear in predicting accuracy. This is because family income tracked variation in language and cognitive abilities in our samples by design, but this also correlated with age-related development. In these analyses, we excluded age and retained family income to isolate individual differences (see Freckleton, 2011, for a discussion of this approach). Data sets and analysis codes are available at <https://osf.io/tfmr3/>. Table 1 provides descriptive statistics of measures.

How do global language and cognitive control vary with SES?

DELV-S responses were scored for total errors and reversed for interpretive ease (higher values reflect greater language knowledge). To compare SES differences across the current study and Leech et al. (2017), we used a linear regression to predict DELV-S based on study and SES (Fig. 1A). This revealed a two-way interaction ($t = 2.02, p < .05$). DELV-S increased with SES in prior work ($t = 2.38, p < .05$), but they were unrelated in the current sample ($t = 0.18, p > .80$). In the Stroop task, responses were coded for naming accuracy and were predicted by trial type and SES through logistic mixed-effects models (Fig. 1B). This revealed a two-way interaction ($z = 3.75, p < .001$). SES was associated with accuracy on incongruent trials ($z = 2.43, p < .05$) but not congruent trials ($z = 0.69, p > .40$). This

¹ Note that the agent-first bias was triggered by different NP1 expressions in the current study versus Leech and colleagues (2017) (i.e., novel words: “The blicket”; known words: “The seal”). To confirm that this did not alter detection of SES differences, we compared effect sizes for NP1 × Construction interactions. This tracked the difficulty of revising the agent-first bias across tasks. Partial eta squared (η_p^2) was in fact higher with novel words (.36; Huang & Arnold, 2016) compared with known words (.13; Leech et al., 2017). Based on the prior SES × Construction interaction ($\eta^2 = .06$), comparable effects of cognitive control should be observable in the current study ($\eta^2 = .16$).

Table 1

Descriptive statistics for measures of global language knowledge (DELV-S scores), cognitive control abilities (naming accuracy on the Stroop task), and syntactic parsing (action accuracy in the Sentence task).

	Mean	SD	Minimum	Maximum	Intraclass correlation
DELV-S	12.5	3.9	4	19	–
<i>Stroop task</i>					
Congruent	81.4%	26.1%	0%	100%	$r = .56, p < .001$
Incongruent	59.6%	33.3%	0%	100%	$r = .63, p < .001$
Cost	21.7%	37.9%	–71.0%	100%	$r = .60, p < .001$
<i>Sentence task</i>					
Active actions	74.6%	19.4%	17.0%	100%	$r = .09, p = .56$
Passive actions	25.7%	27.6%	0%	83.0%	$r = .50, p < .001$

Note. $N = 46$. Intraclass correlations assess the strength of relationships between a child's average performance on first-half trials versus second-half trials. DELV-S, Diagnostic Evaluation of Language Variation–Screening Test.

confirms that in the current sample SES was related to cognitive control abilities but not global language knowledge.

How does linguistic knowledge affect syntactic parsing?

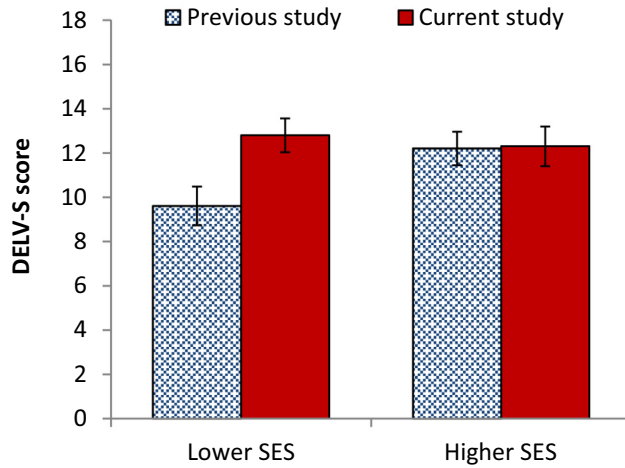
To confirm that sentence properties affected comprehension, we coded actions based on the accuracy of novel word identification (i.e., actives: likely agent; passives: likely theme) and used logistic mixed-effects models to isolate study and construction effects. This revealed a two-way interaction ($z = 4.56, p < .001$). Actives were more accurate than passives, but this difference was greater in the current study ($z = 7.67, p < .001$) compared with prior work ($z = 5.82, p < .01$). Next, we examined how global language affects syntactic revision by predicting accuracy based on construction and DELV-S. This revealed two-way interactions in the current study ($z = 3.84, p < .001$) and prior study ($z = 2.67, p < .01$). We then separated sentences by construction and predicted accuracy based on study and DELV-S. There was no effect for actives ($ps > .20$), but a two-way interaction was found for passives ($z = 2.08, p < .05$). Global language predicted accuracy, but this association was greater in the current study ($z = 2.81, p < .01$) compared with prior work ($z = 2.54, p < .05$). This suggests that linguistic knowledge is associated with syntactic revision beyond general comprehension and that this relationship is present when SES differences are present or absent.

How do global language and cognitive control affect syntactic parsing?

To understand cognitive control effects, we first undertook a conceptual replication of Woodard et al. (2016). A linear regression predicted the average accuracy of passives based on the average accuracy of actives, DELV-S, and Stroop cost (i.e., average accuracy in congruent trials minus incongruent trials; larger values indicate more conflict challenges). Much like correlational analyses, each child contributed a single score per measure. Revising passives was unrelated to actives ($t = 0.09, p > .90$) but increased with higher DELV-S ($t = 2.72, p < .01$) and lower Stroop cost ($t = 2.30, p < .05$). This demonstrates that syntactic revision relates to nonsyntactic conflict when within-individual variance in task performance is minimized.

To isolate cognitive control effects while accounting for within-individual variance, we recruited mixed-effects models to predict comprehension accuracy based on study, SES, and construction (Fig. 2). We found a three-way interaction ($z = 1.99, p < .05$), suggesting that SES effects depend on whether SES tracks variation in global language or cognitive control. When SES differences in global language were present in the prior study, a marginal SES \times Construction interaction was found ($z = 1.89, p < .10$). SES was associated with accuracy for passives ($z = 2.37, p < .05$) but not actives ($z = 1.01, p > .30$). However, when SES tracked cognitive control in the current study, it was associated with improved comprehension ($z = 2.03, p < .05$) with no construction difference ($z = 0.68, p > .40$). This suggests that cognitive control effects might not be specific to syntactic revision. Follow-up analyses estimated Bayes factor (null/alternative) for incongruent Stroop trials and found that the data

A. Global language knowledge



B. Cognitive control abilities

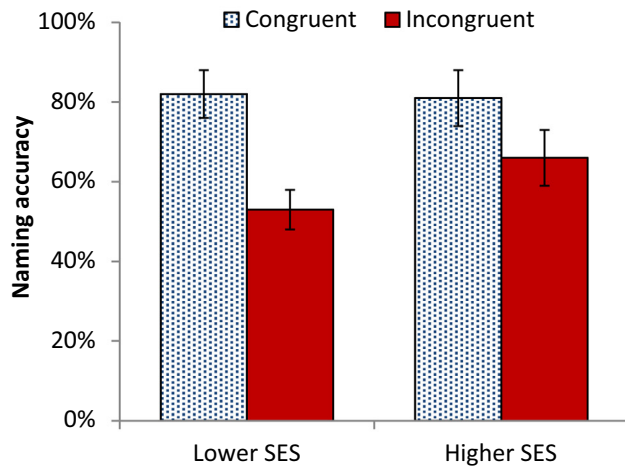
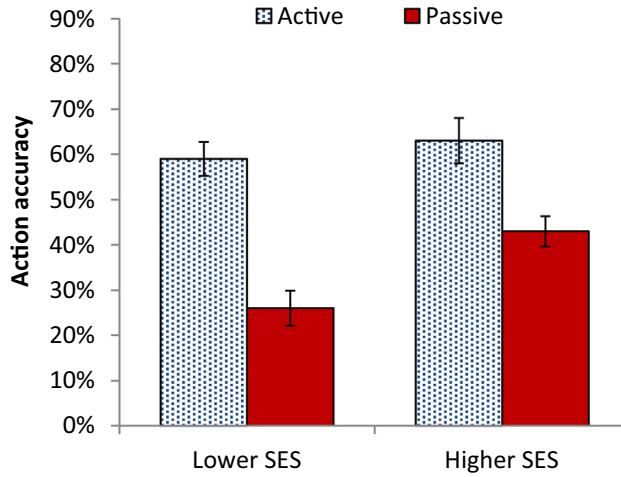


Fig. 1. Across SES background, global language knowledge in the current study ($N = 46$) and the previous study ($N = 60$; Leech et al., 2017) based on the Diagnostic Evaluation of Language Variation–Screening Test (DELV-S) (A) and cognitive control abilities in the current study based on the Stroop task (B). To illustrate SES patterns, samples were split based on median income when this information was available and on school status when it was not.

were 2.66 times more likely to occur under a model that included SES effects compared with a model without them. For passive trials, however, the data were only 0.38 times more likely to occur under a model with SES effects compared with a model without them.²

² We also considered the role of sample size (larger in prior work) and family income (higher in current work) in generating discrepant SES effects. Thus, we matched a subset of the prior sample for sample size and family income to the current sample ($ps > .20$). SES and global language remained correlated ($r = .38$, $p < .05$). Importantly, a marginal two-way interaction between construction and family income was found ($z = 1.93$, $p = .05$). Similar to the full sample, patterns in the subset were driven by SES effects on passives ($z = 2.30$, $p = .02$) but not actives ($z = 1.18$, $p > .20$). This suggests that SES differences in syntactic revision are mediated by relations to global language knowledge and not numerical differences in sample properties.

A. Previous study



B. Current study

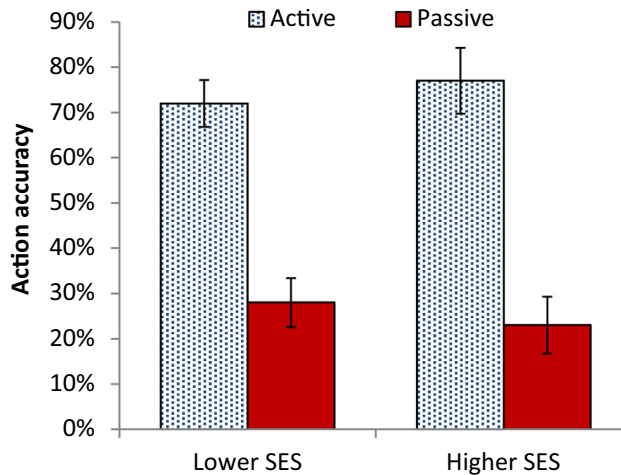


Fig. 2. Across SES background, action accuracy in the Sentence task in the previous study ($N = 60$; Leech et al., 2017) (A) and the current study ($N = 46$) (B) separated by construction. To illustrate SES patterns, samples were split based on median income when this information was available and on school status when it was not.

Discussion

This study examined how global language and cognitive control abilities affect children's recovery from syntactic mispredictions. When SES background tracked variation in global language (Huang et al., 2017; Leech et al., 2017), corresponding effects on revision were present. In contrast, when SES tracked cognitive control but not global language, effects on revision were absent. Although our findings do not rule out an account where "relative underdevelopment [of cognitive-control systems] would result in such an inability to revise initial parsing commitments" (Novick et al., 2005, p. 277), they do suggest that such effects may be intimately tied to developing linguistic knowledge. These

findings also raise questions of how cognitive control effects are instantiated. Do they reflect the consequence of syntactic parsing or altered attention to signal properties? How do experience and maturation influence cognitive control systems across ages, individuals, and tasks? Methodologically, our findings also highlight the challenges of interpreting correlations that do not account for within-individual variability in task performance (e.g., Woodard et al., 2016). Although we addressed this in part through analyses that separately model trial-level performance, more basic questions remain about how variation in experimental tasks maps onto cognitive traits (e.g., Stroop task → cognitive control ability, passives → syntactic revision ability) and whether tasks that are crafted to detect universal processes generate valid measures of individual variation (Hedge et al., 2018).

Theoretically, our findings add to a growing literature revealing language development effects on sentence prediction (Borovsky, Elman, & Fernald, 2012; Mani & Huettig, 2012; Nation, Marshall, & Altmann, 2003; but cf. Gambi, Pickering, & Rabagliati, 2016) and revision (Anderson, Farmer, Goldstein, Schwade, & Spivey, 2011; Huang et al., 2017; but cf. Woodard et al., 2016). They are also consistent with work demonstrating that experience with distributional properties of input (e.g., likelihood of verbs occurring with verb phrase [VP] or NP attachment PPs) influences syntactic parsing in adults and children (Qi, Yuan, & Fisher, 2011; Ryskin, Qi, Duff, & Brown-Schmidt, 2017). Together, this suggests that sentence comprehension is influenced by the acquisition of fine-grained knowledge over one's lifetime. Nevertheless, this raises questions of *why* the kindergarten path effect emerges in the first place. How do developmental limitations in linguistic knowledge influence strategies for interpreting sentences?

One hypothesis is that overreliance on predictions maximizes the likelihood of accurate interpretation. A basic problem of acquisition is that children often lack adequate knowledge to interpret all words in all sentences. Thus, they regularly encounter uncertainty about signal properties (e.g., did speakers say “eaten” or “eating”?). Critically, rational inference models that explain interpretive strategies across challenging communicative contexts (Gibson, Bergen, & Piantadosi, 2013; Levy, Bicknell, Slattery, & Rayner, 2009) may shed light on causes of developmental differences. All listeners must infer speakers' intended meaning from rapidly unfolding cues in speech signals [posterior probability: $p(\text{meaning}_n | \text{cue}_n)$]. Whereas adults can readily infer meanings based on vast knowledge of lexically specific cues [likelihood: $p(\text{cue}_n | \text{meaning}_N)$], children may rely instead on meanings that are shared across most sentences [prior: $p(\text{meaning}_N)$]. An example of this may be the agent-first bias. When children predict meanings through this cue, they may be leveraging their aggregated experience of the word order of sentences to anticipate current meanings. Importantly, this strategy introduces challenges when predictions are wrong. Because speech signals are fleeting and retrieving alternative structures is difficult with limited knowledge, children may often ignore late-arriving conflicts. This suggests that rather than a side effect of cognitive maturation, the kindergarten path effect may reflect interpretive strategies that enable children to understand sentences when signal properties are uncertain.

This hypothesis also provides a more nuanced account of SES differences in comprehension. In prior work (e.g., Leech et al., 2017), one striking pattern was children's proficiency with passives when revision was unnecessary. This demonstrates that infrequent constructions are acquired, even when input is relatively sparse. However, accessing this knowledge during comprehension may critically depend on other input properties. For higher-SES groups, increased quantity may offer ample opportunities to estimate lexically specific patterns. Likewise, greater diversity may enhance the need to access verb-specific properties because idiosyncratic predicates often conflict with canonical frames (e.g., passives, unaccusatives). For lower-SES groups, however, lower input quantity and diversity may make an agent-first bias more informative for predicting meaning because estimates of verb biases are noisy with less input. Moreover, these cues may imply similar meanings in canonical frames. This suggests that across SES children may acquire distinct strategies for interpreting sentences that are likely to occur in their input. Although a rigorous test of this hypothesis awaits future research, this study highlights the need for more detailed descriptions of language acquisition. Because input and outcomes are mediated by children's sentence comprehension, understanding these pathways depends on isolating interpretive strategies both when listeners know little about their language and when they know more.

Acknowledgments

This work was supported by grants from the UMD ADVANCE Program for Inclusive Excellence to YTH and the Mary Cobey Martin Fund for Research Excellence to EH. This work benefited from conversations with Meredith Rowe, Jared Novick, Rachel Ryskin, and Sarah Brown-Schmidt. We thank the families and administrators at Meade Village Head Start (Anne Arundel County, MD) for taking part in this study.

References

- Anderson, S. E., Farmer, T. A., Goldstein, M., Schwade, J., & Spivey, M. (2011). Individual differences in measures of linguistic experience account for variability in the sentence processing skill of five-year-olds. In I. Aron & E. V. Clark (Eds.), *Experience, variation, and generalization: Learning a first language (Trends in Language Acquisition Research)* (Vol. 7, pp. 203–221). Philadelphia: John Benjamins.
- Barr, D., Levy, R., Scheepers, C., & Tily, H. (2013). Random effects structure for confirmatory hypothesis testing: Keep it maximal. *Journal of Memory and Language*, *68*, 255–278.
- Bates, D., Maechler, M., Bolker, B., & Walker, S. (2015). lme4: Linear mixed-effects models using Eigen and S4. R package version 1.1-10. <http://CRAN.R-project.org/package=lme4>.
- Beveridge, M., Jarrold, C., & Pettit, E. (2002). An experimental approach to executive finger-printing in young children. *Infant and Child Development*, *11*, 107–123.
- Borovsky, A., Elman, J. L., & Fernald, A. (2012). Knowing a lot for one's age: Vocabulary skill and not age is associated with anticipatory incremental sentence interpretation in children and adults. *Journal of Experimental Child Psychology*, *112*, 417–436.
- Choi, Y., & Trueswell, J. C. (2010). Children's (in) ability to recover from garden paths in a verb-final language: Evidence for developing control in sentence processing. *J. Exp. Child Psychol.*, *106*(1), 41–61.
- Freckleton, R. P. (2011). Dealing with collinearity in behavioural and ecological data: Model averaging and the problems of measurement error. *Behavioral Ecology and Sociobiology*, *65*, 91–101.
- Gambi, C., Pickering, M. J., & Rabagliati, H. (2016). Beyond associations: Sensitivity to structure in pre-schoolers' linguistic predictions. *Cognition*, *157*, 340–351.
- Gibson, E., Bergen, L., & Piantadosi, S. (2013). The rational integration of noisy evidence and prior semantic expectations in sentence interpretation. *Proceedings of the National Academy of Sciences of the United States of America*, *110*, 8051–8056.
- Hart, B., & Risley, T. (1995). *Meaningful differences in everyday experience of young American children*. Baltimore, MD: Paul Brookes.
- Hedge, C., Powell, G., & Sumner, P. (2018). The reliability paradox: Why robust cognitive tasks do not produce reliable individual differences. *Behavior Research Methods*, *50*, 1166–1186.
- Hoff, E. (2003). The specificity of environmental influence: Socioeconomic status affects early vocabulary development via maternal speech. *Child Development*, *74*, 1368–1378.
- Huang, Y., & Arnold, A. (2016). Word learning in linguistic context: Processing and memory effects. *Cognition*, *157*, 71–87.
- Huang, Y., Leech, K., & Rowe, M. R. (2017). Exploring socioeconomic differences in syntactic development through the lens of real-time processing. *Cognition*, *159*, 61–75.
- Huang, Y., Zheng, X., Meng, X., & Snedeker, J. (2013). Children's assignment of grammatical roles in the online processing of Mandarin passive sentences. *Journal of Memory and Language*, *69*, 589–606.
- Hurewitz, F., Brown-Schmidt, S., Thorpe, K., Gleitman, L., & Trueswell, J. (2000). One frog, two frog, red frog, blue frog: Factors affecting children's syntactic choices in production and comprehension. *Journal of Psycholinguistic Research*, *29*, 597–626.
- Kidd, E., Stewart, A. J., & Serratrice, L. (2011). Children do not overcome lexical biases where adults do: The role of the referential scene in garden-path recovery. *Journal of Child Language*, *38*, 222–234.
- Leech, K., Rowe, M., & Huang, Y. (2017). Variations in the recruitment of syntactic knowledge contribute to SES differences in syntactic development. *Journal of Child Language*, *44*, 995–1009.
- Levy, R., Bicknell, K., Slattery, T., & Rayner, K. (2009). Eye movement evidence that readers maintain and act on uncertainty about past linguistic input. *Proceedings of the National Academy of Sciences of the United States of America*, *106*, 21086–21090.
- Mani, N., & Huettig, F. (2012). Prediction during language processing is a piece of cake—But only for skilled producers. *Journal of Experimental Psychology: Human Perception and Performance*, *38*, 843–847.
- Mazuka, R., Jincho, N., & Onishi, H. (2009). Development of executive control and language processing. *Language and Linguistics Compass*, *3*, 59–89.
- Nation, K., Marshall, C. M., & Altmann, G. T. M. (2003). Investigating individual differences in children's real-time sentence comprehension using language-mediated eye movements. *Journal of Experimental Child Psychology*, *86*, 314–329.
- Noble, K. G., McCandliss, B. D., & Farah, M. J. (2007). Socioeconomic gradients predict individual differences in neurocognitive abilities. *Developmental Science*, *10*, 464–480.
- Noble, K. G., Norman, M. F., & Farah, M. J. (2005). Neurocognitive correlates of socioeconomic status in kindergarten children. *Developmental Science*, *8*, 74–87.
- Noble, K. G., Wolmetz, M. E., Ochs, L. G., Farah, M. J., & McCandliss, B. D. (2006). Brain-behavior relationships in reading acquisition are modulated by socioeconomic factors. *Developmental Science*, *9*, 642–654.
- Novick, J. M., Trueswell, J. C., & Thompson-Schill, S. L. (2005). Cognitive control and parsing: Reexamining the role of Broca's area in sentence comprehension. *Cognitive, Affective, and Behavioral Neuroscience*, *5*, 263–281.
- Omaki, A., Davidson White, I., Goro, T., Lidz, J., & Phillips, C. (2014). No fear of commitment: Children's incremental interpretation in English and Japanese wh- questions. *Language Learning and Development*, *10*, 206–233.

- Qi, Z., Fisher, C., & Brown-Schmidt, S. (2011). On recovering from “kindergarten path” errors: Referential context and executive functioning influence children’s online ambiguity resolution. Paper presented at the 17th Annual Conference on Architectures and Mechanisms for Language Processing (AMLaP), Paris.
- Qi, Z., Yuan, S., & Fisher, C. (2011). Where does verb bias come from? Experience with particular verbs affects online sentence processing. In N. Danis, K. Mesh, & H. Sung (Eds.), *Proceedings of the 35th annual Boston University Conference on Language Development* (Vol. 35, pp. 500–512). Somerville, MA: Cascadilla Press.
- Rowe, M. L. (2008). Child-directed speech: Relation to socioeconomic status, knowledge of child development and child vocabulary skill. *Journal of Child Language*, 35, 185–205.
- Ryskin, R. A., Qi, Z., Duff, M. C., & Brown-Schmidt, S. (2017). Verb biases are shaped through lifelong learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 43, 781–794.
- Seymour, H. N., Roeper, T., & De Villiers, J. G. (2003). *DELV: Diagnostic evaluation of language variation: Screening test*. San Antonio, TX: PsychCorp.
- Trueswell, J. C., Sekerina, I., Hill, N. M., & Logrip, M. L. (1999). The kindergarten-path effect: Studying on-line sentence processing in young children. *Cognition*, 73, 89–134.
- Woodard, K., Pozzan, L., & Trueswell, J. C. (2016). Taking your own path: Individual differences in executive function and language processing skills in child learners. *Journal of Experimental Child Psychology*, 141, 187–209.