COGNITIVE SCIENCE
A Multidisciplinary Journal

Cognitive Science 46 (2022) e13076

ISSN: 1551-6709 online
DOI: 10.1111/cogs.13076

Who "it" is influences what "it" does: Discourse effects on children's syntactic parsing

Yi Ting Huang,[a,b] Zoe Ovans[a,b]

[a]*Department of Hearing and Speech Sciences, University of Maryland College Park*
[b]*Neuroscience and Cognitive Science Program, University of Maryland College Park*

Received 13 March 2020; received in revised form 30 October 2021; accepted 12 November 2021

Abstract

Children often interpret first noun phrases (NP1s) as agents, which improves comprehension of actives but hinders passives. While children sometimes withhold the agent-first bias, the reasons remain unclear. The current study tests the hypothesis that children default to the agent-first bias as a "best guess" of role assignment when they face uncertainty about sentence properties. Thus, rather than always relying on early-arriving cues, children can attend to different sentence cues across communicative contexts. To test this account, we manipulated interpretive uncertainty by varying cues to the discourse status of initial arguments (referring to new vs. given entities) and measured interpretation accuracy for active (where the agent-first bias predicted verb morphology) and passive sentences (where the two conflicted). Across three experiments, we found that children relied on the agent-first bias more when new discourse entities were signaled by definite NP1s, reference to unmentioned entities, and novel words. This, in turn, led to higher accuracy for actives relative to passives. In contrast, when given entities were implied through pronoun NP1s, reference to mentioned entities, and known words, children avoided the agent-first bias and instead assigned roles using more reliable but later-arriving verb morphology. This led to similar comprehension accuracy across constructions. These findings suggest that children simultaneously interpret relations between sentences (e.g., discourse continuity) and within sentences (e.g., role assignment), such that commitments to the former can influence parsing cues for the latter.

Keywords: Syntactic parsing; Agent-first bias; Noisy channel; Passives; Discourse continuity

Correspondence should be sent to Yi Ting Huang, Department of Hearing and Speech Sciences, University of Maryland College Park, 0100 Samuel J. LeFrak Hall, College Park, MD 20742 USA. E-mail: ythuang1@umd.edu

1. Introduction

During spoken-language comprehension, children link current sentence cues to existing linguistic knowledge to infer a speaker's intended meaning. A large body of work has shown that during this process, toddlers and school-aged children alike privilege early-arriving cues in sentences over late-arriving ones, a pattern dubbed the "kindergarten-path effect" (Choi & Trueswell, 2010; Huang, Zheng, Meng, & Snedeker, 2013; Kidd, Stewart, & Serratrice, 2011; Lidz, White, & Baier, 2017; Omaki, Davidson White, Goro, Lidz, & Phillips, 2014; Trueswell, Sekerina, Hill, & Logrip, 1999; Woodard, Pozzan, & Trueswell, 2016). For example, in sentences like (1), children incrementally interpret first noun phrases (NP1s) as agents (i.e., they used an "agent-first bias"). This enhances comprehension of actives where *the seal* is the agent, but hinders passives which require revision after the past participle reveals *the seal* is instead the patient (Huang & Hollister, 2019; Huang, Leech, & Rowe, 2017; Abbott-Smith, Chang, Rowland, Ferguson, Pine, 2017; Deen et al., 2018).

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

Developmental differences in sentence comprehension are well studied, but less is known about how communicative contexts influence parsing dynamics during acquisition. Do children always prefer early-arriving cues over later-arriving ones, or do parsing strategies vary with context-specific properties? Several studies show that children ignore discourse context during parsing (Snedeker & Trueswell, 2004; Trueswell et al., 1999; Kidd & Bavin, 2005; Choi & Trueswell, 2010; Weighall, 2008; Woodard et al., 2016), but leave open the possibility that developmental parsing is affected by other discourse cues, such as the givenness of particular referents. Such evidence would be consistent with context-dependent parsing in adults (Ferreira, Ferraro, & Bailey, 2002; Gibson, Bergen, & Piantadosi, 2013; Levy, Bicknell, Slattery, & Rayner, 2009), and would suggest that children attend to different sentence cues based on properties of the context, rather than applying form-to-meaning relations deterministically across sentences. This probabilistic strategy for parsing may be useful during development by enabling young listeners to optimize how they calculate utterance meaning with available linguistic knowledge.

We turn to the agent-first bias as an example of a heuristic that children may rely on when faced with uncertainty. This bias is based on broad-scale regularities of English word order (e.g., Chan, Lieven, & Tomasello, 2009), making this particular form-to-meaning relations easy to infer during development. We examine how this bias varies with a systematic source of uncertainty during sentence comprehension, namely whether NP1s refer to new or given entities in the discourse. In the remainder of the Introduction, we summarize prior work on the developmental origins of the agent-first bias, and introduce evidence that children withhold this bias under specific contexts (Huang et al., 2013). Next, we consider potential explanations for this pattern, and outline three experiments that manipulate NP1 properties and measure comprehension of active and passive sentences. Based on these findings, we suggest that the decisions underlying moment-to-moment processing and the forces that guide which linguistic phenomena children acquire are distinct phenomena (akin to how day-to-day weather and long-term climate are related but separable constructs), and propose that accurate estimates of these time scales will improve our understanding of the factors influencing language acquisition.

1.1. The origin and development of the agent-first bias

The agent-first bias is an incredibly well-studied phenomenon in language acquisition. Bever (1970) observed a curious pattern whereby comprehension of passives improves from age 2 to 6, but unexpectedly dips at age 4. This age coincides with improved accuracy with implausible actives (e.g., *The dog pats the mother*), and Bever argued that this U-shape function reflected developmental changes in comprehension strategies. While 2-year-olds calculate sentence meanings via lexical semantics (e.g., a word list like *dog*, *pat*, *mother* implies that *The mother pats the dog*), 4-year-olds adopt structurally guided constraints like the agent-first bias. Likewise, De Villiers and De Villiers (1973) found that when lexical semantic cues are absent, all 2-year-olds correctly interpret reversible actives (e.g., *The girl is kicking the boy*), and less advanced 2-year-olds produce mixed responses for passives (e.g., *The girl is kicked by the boy*). However, more advanced 2-year-olds adopt an agent-first bias and generate widespread role-reversal errors. This suggests that the agent-first bias arises during development, and its use during comprehension can limit interpretation of specific constructions.

A decade later, Bates and MacWhinney's Competition Model (1981) reframed the agent-first bias as one of many parsing strategies resulting from inferences of form-to-meaning relations during development. Children generate interpretations based on sentence cues that reliably predict correct argument roles and occur often in the input. Cross-linguistic comparisons reveal different parsing strategies that can arise from diverging input statistics (MacWhinney, Bates, & Kliegl, 1984; MacWhinney, Pléh, & Bates, 1985). For example, English's subject-verb-object (SVO) order leads to many sentences where NP1s are agents (Chan et al., 2009), relative to pro-drop languages like Italian, which often omits salient subjects in the discourse (Bates, McNew, MacWhinney, Devescovi, & Smith, 1982). Thus, when identifying agents in sentences, English-speaking 2-year-olds isolate NP1s, while Italian learners focus on animate NPs (Bates et al., 1984). Likewise, German (e.g., scrambling) and Cantonese (e.g., pro-drop) permit flexible word order, leading to a weaker agent-first bias in 2- and 3-year-old learners compared to English-speaking counterparts (Chan et al., 2009). This suggests that children are sensitive to broad-scale regularities in how languages assign grammatical roles by age 3, and this influences the extent to which they employ the agent-first bias.

Finally, Chang, Dell, and Bock (2006)'s dual-path model formalized how developmental changes in input alter the agent-first bias in English, and applied this to understanding comprehension patterns across constructions. For example, 2-year-olds correctly interpret active transitives (e.g., *She is glorping the spoon* → infer that NP2 is a patient) but miss-assign roles for *with*-intransitives (e.g., *She is glorping with the spoon* →NP2 is an instrument, misconstrued as a patient) (Hirsh-Pasek & Golinkoff, 1996; Lidz et al., 2017). These errors can be explained through a connectionist architecture that encodes lexical concepts and grammatical roles of prior words to predict up-coming ones. Correct predictions strengthen connections between cues and roles while incorrect predictions re-weight mappings through error-based learning. Since subject NPs are often causative agents in English (e.g., transitives, datives, locative alternations, benefactives), the model readily infers an agent-first bias. This facilitates comprehension of active transitives but produces errors with *with*-intransitives. As experience with alternative constructions (e.g., intransitives, passives) accumulates with age, the model increases its reliance on postverbal cues, and this decreases misinterpretations. Thus, the dual-path model provides a useful framework for understanding developmental changes in parsing strategies.

1.2. Agent-first bias as a heuristic to overcome contextual uncertainty

While past research offers detailed descriptions of the development of the agent-first bias, less is known about *where* and *why* this heuristic is useful for comprehension. In fact, such questions are difficult to pose within theories that treat ontogenetic and chronometric procedures as the same (e.g., competition model, dual-path model), a point we will return to in Section 5. We suggest that treating development and processing as separable constructs enables specific hypotheses about how parsing strategies may vary across communicative contexts during language acquisition. In the case of the agent-first bias, we propose that children default to this "best guess" of role assignment when they encounter uncertainty about sentence properties. Thus, this heuristic is informed by linguistic knowledge but not identical to it.

Our hypothesis is inspired by psycholinguistic accounts of variable parsing strategies in adults, for example, Good-Enough Processing (Ferreira, & Patson, 2007; Ferreira, Bailey, & Ferraro, 2002) and Noisy-Channel Parsing (Gibson et al., 2013; Levy et al., 2009). Adults misinterpret implausible passives using an agent-first bias (e.g., *The girl was kicked by the ball* → NP1 misconstrued as the agent), particularly when sentences are interleaved with anomalous forms (Ferreira, 2003; Gibson et al., 2013; MacWhinney et al., 1984). This suggests that uncertainty about signal properties during communication may enhance reliance on canonical parsing cues. From a Bayesian perspective, this tradeoff makes sense. Since adults infer sentence meanings by combining expectations of what is likely to be said (prior) with what is actually said (likelihood), they maximize accurate interpretations by discounting evidence when signal properties are ambiguous. Importantly, this same strategy may be useful during acquisition, when limited knowledge of words and rules introduces ample uncertainty about signal properties. Heuristics like the agent-first bias may enable children to interpret sentences using knowledge of broad-scale regularities in language when finer-grained ones have yet to be fully mastered.

To the best of our knowledge, no research has systematically tested this hypothesis in children, but two studies provide tentative support. First, like English, the SVO word order in French generates an agent-first bias. However, French also features a right-dislocation construction, which highlights topical referents through a prosodic break (e.g., *He$_i$ ate (,) the rabbit$_i$* → NP2 is the agent). Two-year-olds recruit this cue to correctly interpret NP2s as subjects of intransitives (Dautriche et al., 2014). However, when novel verbs increase uncertainty, children ignore prosody and misapply the agent-first bias (e.g., *He$_i$ dase (,) the rabbit$_i$* → NP2 misconstrued as the patient). Similarly, when English-speaking 5-year-olds recruit the agent-first bias, this leads to accurate comprehension of actives but difficulties with passives (Huang & Arnold, 2016). Importantly, when speech signals are acoustically distorted to simulate listening through cochlear implants, accuracy remains high with actives but worsens with passives (Martin, Goupell, & Huang, in press). Together, this suggests that diverse aspects of communicative contexts can generate interpretive uncertainty (e.g., unfamiliar words, acoustic degradation), which in turn can increase reliance on canonical parsing cues.

Building on these findings, the current study systematically varies another source of uncertainty during comprehension, namely whether NP1s are new or given entities in the discourse. Across development, children from a young age expect current and prior sentences to relate (Branigan & Messenger, 2016; Peter, Chang, Pine, Blything, & Rowland, 2015), and exploit linguistic cues to discourse continuity when interpreting pronouns (Hartshorne, Nappa, & Snedeker, 2015; Pyykkönen, Matthews, & Järvikivi, 2010; Song & Fisher, 2005; Arnold, Castro-Schilo, Zerkle, & Rao, 2019), prosody (Arnold, 2008; Ito, Bibyk, Wagner, & Speer, 2014), and novel words (Fisher, Jin, & Scott, 2019; Horowitz & Frank, 2015; Sullivan & Barner, 2016). Critically, cues to discourse status may also impact strategies for role assignment. Since spoken sentences unfold incrementally, children face decisions about argument interpretation before hearing all relevant parsing cues. Thus, we propose that linguistic cues that imply previously unmentioned (new) discourse entities may enhance interpretive uncertainty and reliance on the agent-first bias, while cues that imply previously mentioned (given) entities may reduce uncertainty and the agent-first bias. Delaying role assignment may benefit comprehension by allowing children to assign roles using more reliable cues that occur later in sentences (e.g., verb-specific semantics/syntax).

1.3. Test case: Interpretation of active and passive sentences

Discourse effects may explain paradoxical patterns within children's interpretations of actives and passives (Huang et al., 2013). Similar to English, Mandarin's SVO word order generates an agent-first bias. Thus, when Mandarin-speaking 5-year-olds hear sentences with definite NP1s like (2), they are less accurate at interpreting passives (i.e., BEI) compared to actives (i.e., BA). Curiously, however, when Mandarin-speaking children encounter sentences with pronoun NP1s like (3), they now interpret passives as accurately as actives. By reversing the order of definite NPs and pronoun (e.g., *the seal* vs. *it*), NP1s differ in their discourse implications. Definite NPs often signal new entities (Arnold, 1998; Ariel, 1988; Chafe, 1987; Gundel, Hedberg, & Zacharski, 1993; Rohde & Frank, 2014), leading to greater uncertainty in sentence-initial position. This, in turn, may trigger an agent-first bias that improves actives but introduces conflicts for passives. In contrast, when pronoun NP1s are associated with less uncertainty, this may avert the agent-first bias, and enable children to interpret actives (i.e., BA) and passives (i.e., BEI) on equal footing. Consistent with this hypothesis, children generate correct fixations to pronoun referents in sentences like (3) after NP2 onset. This suggests

Table 1
Predictions for the agent-first bias across manipulations of NP1 status. For each account, listed condition(s) are ones where comprehension accuracy is expected to be lower for passives compared to actives

ACCOUNTS	EXPERIMENT 1 (Definite vs. Pronoun NP1)	EXPERIMENT 2 (Conjoined vs. Single-NP prime)	EXPERIMENT 3 (Novel vs. Known NP1)
Discourse status *(agent-first bias stronger with new discourse entities)*	Definite NP1	Conjoined-NP	Novel NP1
Language-specific preference *(pronoun NPs are dispreferred in English)*	Pronoun NP1	Low accuracy for all conditions	High accuracy for all conditions
Noun-phrase frequency *(infrequent NPs are harder to parse)*	Definite NP1	High accuracy for all conditions	Low accuracy for all conditions
Referential ambiguity *(agent-first bias stronger with reference restriction)*	Definite NP1	Single-NP	Known NP1

that when agent-first predictions are avoided, children exploit late-arriving cues to assign NP2 roles and identify NP1s on this basis.

(2)	a. Active:	海 豹 把 它 很 快 就 吃 掉 了	Seal BA it quickly eat. (Translation: *The seal is quickly eating it*)
	b. Passive:	海 豹 被 它 很 快 就 吃 掉 了	Seal BEI it quickly eat. (Translation: *The seal is quickly eaten by it*)
(3)	a. Active:	它 把 海 豹 很 快 就 吃 掉 了	It BA seal quickly eat. (Translation: *It is quickly eating the seal*)
	b. Passive:	它 被 海 豹 很 快 就 吃 掉 了	It BEI seal quickly eat. (Translation: *It is quickly eaten by the seal*)

At the discourse level, the notion of uncertainty we wish to invoke is that referring to new entities incurs a cost compared to given entities since they require activating previously inactive referents (e.g., Arnold, 2010). This relative unpredictability may promote incremental role assignment through canonical parsing cues. However, definite and pronoun NPs vary along multiple dimensions. This introduces alternative explanations of interactions between NP1 expression and construction. In the current study, we evaluate the Discourse-status account against three potential explanations of Huang and colleague (2013) patterns. These hypotheses are inspired by prior discussions of developmental processing in the literature (Table 1).

Language-specific preference: Since Huang et al. (2013) examined comprehension in Mandarin, it is possible that the pronoun-NP1s advantage reflects a language-specific preference for sparse sentential contexts. In transitive sentences, English-speaking 2-year-olds are more likely to learn novel verbs with definite NPs (e.g., *The man is pilking the balloon* → learn that *pilking* refers to causative event) compared to pronoun NPs (e.g., *He is pilking it* → no verb learning) (Arunachalam & Waxman, 2011, 2015; Waxman, Lidz, Braun, & Lavin, 2009). However, in pro-drop languages like Korean (Arunachalam, Leddon, Song, Lee, & Waxman, 2013) and Japanese (Imai et al., 2008; Imai, Haryu, & Okada, 2005), children show *less* learning with definite NPs compared to omitted arguments. Since Mandarin is also a pro-drop language, it is possible that prior difficulties with definite NPs reflect a preference to identify salient referents with reduced expressions. To test this hypothesis, the current study examines sentences like (2) and (3) in English, a nonpro-drop language (Experiment 1). If prior NP1 effects reflect language-specific preferences, then English-speaking children may be more accurate with semantically rich definite NPs. If prior effects reflect relations between NP1 expressions and discourse status, then English patterns should mirror ones found in Mandarin.

NP frequency: Prior advantages with pronoun NP1s may reflect the general ease of interpreting frequent forms, rather than relations to the agent-first bias. In intransitive sentences, English-speaking 2- and 3-year-olds are more likely to learn novel verbs with pronoun NPs (e.g., *It is blicking* → *blicking* refers to self-caused motion) compared to definite NPs (e.g., *The flower is blicking* →no verb learning) (He, Kon, & Arunachalam, 2020; Lidz, Bunger, Leddon, Baier, & Waxman, 2010), and repeated mention of definite NPs improves learning (e.g., *Do you see the flower? There's the flower again. It's a flower. The flower is blicking* → verb learning). When subject NPs are easier to retrieve from the lexicon, this enhances parsing and promotes word learning. To test this hypothesis, the current study examines the same pronoun NP1s in sentences like (3), when they refer to new or given referents (Experiment 2). If comprehension is related to argument retrieval from the lexicon, parsing may be similar when sentences are identical. If the agent-first bias is adopted due to uncertainty, the same sentence may generate different interpretations depending on the discourse status of NP1.

Referential ambiguity: Huang et al. (2013) originally attributed increased success with passives to the referential ambiguity of pronoun NP1s. During the task, spoken sentences were presented with displays featuring thematically related objects (e.g., seal, fish, shark), and the authors noted that unlike definite NP1s, pronoun NP1s initially cannot be assigned to a referent. By blocking reference restriction, prior work may have unnaturally postponed role assignment until children heard the active or passive markers. This suggests that the agent-first bias may in fact be children's default strategy for interpreting sentences, but this pathway is short circuited when they encounter uninterpretable expressions (e.g., pronoun NPs without antecedents). To test this hypothesis, the current study examines sentences where definite NP1s refer to novel entities, for example, *The blicket* (Experiment 3). If prior NP1 effects reflect referential ambiguity, then novel expressions should pattern like pronoun NP1s and decrease the agent-first bias. In contrast, if prior effects reflect relations to discourse status, then novel expressions should increase interpretative uncertainty, and enhance reliance on the heuristic.

In three experiments, we manipulate the strength of the agent-first bias (weak vs. strong) and construction (active vs. passive), and examine English-speaking children's comprehension in an act-out task. Critical sentences vary. Similar to Huang et al. (2013), sentences are paired with displays that feature a mentioned object (e.g., seal), likely agent (e.g., shark), and likely theme (e.g., fish). Children's selection of the likely agent or theme provides a measure of their role assignment. Across experiments, the discourse-status account predicts that reference to given entities should reduce interpretive uncertainty and the agent-first bias. This, in turn, would enable role assignment based on verb morphology for actives (e.g., *eating*) and passives (e.g., *eaten*) and lead to comparable accuracy across constructions. Conversely, reference to new discourse entities should increase uncertainty and the agent-first bias. This, in turn, would facilitate interpretations for actives, but create kindergarten-path difficulties for passives.

2. Experiment 1

To examine whether the patterns found by Huang et al. (2013) are specific to pro-drop languages, we modified their materials for testing in English. In actives like (4a), definite NP1s are agents (e.g., *the seal* is the predator) while pronoun NP2s are themes (e.g., *it* = a fish). In passives like (4b), definite NP1s are themes (e.g., *the seal* is the prey) while pronoun NP2s are agents (e.g., *it* is a shark). These patterns reverse in sentences like (5). Here, pronoun NP1s are agents in actives (e.g., *it* is a shark) and themes in passives (e.g., *it* is a fish). Our assumption is that discourse status is inferred through probabilistic relations to definite and pronoun expressions in input statistics (Arnold, 1998; Ariel, 1988; Rohde & Frank, 2014), and that adults and children draw on this knowledge when interpreting sentences (Warren & Gibson, 2002; Brandt, Kidd, Lieven, & Tomasello, 2009; Gordon, Hendrick, & Johnson, 2004).

(4)	Definite NP1	a. Active:	The seal is quickly eating it.
		b. Passive:	The seal is quickly eaten by it.
(5)	Pronoun NP1	a. Active:	It is quickly eating the seal.
		b. Passive:	It is quickly eaten by the seal.

Recall that the language-specific account suggests that semantically rich arguments offer parsing advantages in English (Arunachalam & Waxman, 2011, 2015; Waxman et al., 2009). This predicts that unlike Mandarin, difficulties with passives will now emerge in sentences with pronoun NP1s. In contrast, the discourse-status account predicts that construction differences are mediated by relationships between uncertainty and the agent-first bias. Similar to Mandarin, passives should be less accurate compared to actives when definite NP1s signal new entities (which enhances the bias), and constructions should be similar when pronoun NP1s signal given entities (which weakens the bias). Note that the frequency and referential-ambiguity accounts predict identical patterns but for different reasons. The frequency account suggests that definite NP1s are harder to retrieve from memory, and this hinders integration into infrequent structures like passives. The referential-ambiguity account suggests that the agent-first bias requires reference restriction, which is only possible with definite NP1s.

2.1. Methods

2.1.1. Participants

Forty English-speaking children (20 male, 20 female) were recruited from private schools in the Washington, D.C. metro area. This sample size was based on prior work in the literature that used act-out tasks to assess sentence comprehension (e.g., Huang et al., 2013; Snedeker & Trueswell, 2004; Trueswell et al., 1999). The mean age was 5.3 ($SD = 0.7$, range = 4.6–7.0). Similar to Huang et al. (2013), NP1 status was manipulated between subjects to avoid perseveration across trials. Half the children were randomly assigned to the definite NP1 condition (sentences like (4)), and the other half to the pronoun NP1 condition (sentences like (5)). There was no significant difference in age (in months) or gender across the two groups ($p > .80$).

2.1.2. Procedure

Children sat in front of a display that was divided into four quadrants. The experimenter labeled objects in each set individually as they were placed on the shelf in a pre-specified order. Each object set was paired with two sentence trials. On each trial, children heard a sentence describing an event. Children were encouraged to pick up the objects and use them to act-out what was said. Once they did this, the trial ended and the objects were returned to their original locations. This was followed by a second sentence describing another event involving the same objects. Once children performed this action, the objects were removed, and the next trial began with a new set of objects. Critical trials were always the first sentence in an object set, while filler trials were either the first or second sentence.

2.1.3. Materials

Three-object sets were created based on a mentioned object (e.g., seal), likely agent (e.g., shark, which is likely to act on the mentioned object), and likely theme (e.g., fish, which the mentioned object is likely to act on). Across items, the plausibility of role relationships was confirmed through norming judgments from adults (i.e., *Based on your knowledge, how likely is X to act on Y?*). The norming data can be found at osf.io/8shgf/. The relative object size provided another cue to role relationships. Likely agents were larger than mentioned objects, which in turn were larger than likely themes. Across trials, the location of the object types varied to ensure that role assignments could not be predicted based on display configuration.

Critical sentences included definite and pronoun NPs but varied the order of occurrence, see (4) and (5). NP1 status contrasted definite NPs (e.g., *the seal*) versus pronoun (i.e., *it*) in sentence-initial position. Construction type varied actives versus passives, which differed in verb morphology (present progressive: *eat<u>ing</u>* vs. past participle: *eat<u>en</u>*) and the presence of the *by*-phrase. This was manipulated within subjects. Across sentences, the *be*-auxiliary and an adverb (e.g., *is quickly*) were embedded between NP1 and the main verb to create an extended period where role assignments could not be informed by event semantics. Twelve critical trials were randomized with 32 filler trials, which were included to divert attention away from manipulated variables without systematically biasing roles. These included agent/theme intransitives (e.g., *sing*, *break*), experiencer-stimulus verbs (e.g., *like*, *scare*), and symmetric predicates (e.g., *dance*, *fight*). Across all experiments, four versions of critical items were used to create four presentation lists, such that each list contained six items in a condition and each item appeared once in every list. All sentences were pre-recorded by a female actor who spoke in a slow and consistent manner. See Appendix A for a full list of critical items.

2.1.4. Coding

Approximately 2.3% of trials were excluded because of experimenter error or because no action was produced. For the remaining trials, a trained research assistant coded videotapes of actions and categorized responses based on the trial condition. *Correct actions* depicted correct role assignments between the mentioned object and a plausible object. For definite NP1/passive and pronoun NP1/active trials, these involved the likely agent doing something to the mentioned object (e.g., making the shark eat the seal). For definite NP1/active and pronoun NP1/ passive trials, these involved the mentioned object doing something to the likely theme (e.g., making the seal eat the fish). *Reverse actions* were those that depicted incorrect role assignments between the mentioned object and a plausible object. For definite NP1/passive and pronoun NP1/active trials, these involved the mentioned object doing something to the likely theme. For definite NP1/active and pronoun NP1/passive trials, these involved the likely agent doing something to the mentioned object. *Ambiguous actions* were those where only the mentioned object was selected or all three objects were selected. A second research assistant confirmed the original coding on 94.8% of trials. Any disagreements between the original and reliability coding were resolved by a third independent coder.

2.2. Results

Since trial-level performance was coded as a series of categorical variables (e.g., correct vs. incorrect actions), we analyzed the data using logistic mixed-effects models (Jaeger, 2008). For each experiment, three models were constructed for each dependent variable: (A) a primary model predicting accuracy (1 vs. 0), (B) a follow-up model predicting reversal errors (1 vs. 0), and (C) a follow-up model predicting ambiguous errors (1 vs. 0). In each case, NP1 status (definite vs. pronoun NP1) and construction (active vs. passive) were modeled as fixed-effects variables, and likelihood ratio tests computed p-values by comparing simpler and more complex models. Effects coding within fixed effects compared condition means to the grand mean, and follow-up analyses separated models separated by one fixed effect to compare within the other. We constructed additional models that included age (in months) and gender as predictors. Since these factors never significantly improved model fit (all p's $>$ .20), we present results from the simpler models. Analyses were implemented using the lme4 software package in R (Bates, Maechler, Bolker, & Walker, 2015).

Following psycholinguistic conventions, we started with maximal models that included both random slopes and intercepts for subjects and items (Barr, Levy, Scheepers, & Tily, 2013), but found that these models often failed to converge. This may reflect a confluence of factors relevant to the current study, including: (1) binary responses (0 or 1) provide limited information about probabilities of observations (Eager & Roy, 2017), (2) item variance is typically small in experimental designs (Jaeger, 2009), and (3) nonconverge often arises when there is insufficient variance in both the random intercept and slope (Moineddin, Matheson, & Glazier, 2007). In response, we adopted an analytical strategy whereby random-effects structures were included when justified by model comparison and supported by the data (Baayen, Davidson, & Bates, 2008; Gries, 2015). In Experiments 1 and 3, this often meant omitting random slopes or intercepts for items, which had far less variance than subjects. In Experiment 2, however, item variance was on par with subject variance since parsing of critical sentences depended on probabilistic interpretation of preceding prime sentences. Here random effects for items were maintained for these models. Summaries of final models are provided in Tables2–6. The data and analysis code can be found at osf.io/8shgf/.

We first examine the accuracy of children's interpretation. Final models included random slopes and intercepts for subjects but omitted them for items. Model comparison revealed no significant difference in the variance explained by models that included items as random effects (p's $> .40$). Fig. 1 illustrates that children were less accurate at interpreting passives ($M = 44\%$, $SD = 37\%$) compared to actives ($M = 68\%$, $SD = 23\%$) in sentences with definite NP1s ($z = 2.05$, $SE = 0.35$, $p < .05$), but this difference was not present with pronoun NP1s ($z = 0.36$, $SE = 0.25$, $p > .70$). Analyses by construction provide additional support that NP1 status influenced comprehension. While passives were marginally less accurate with definite NP1s compared to pronoun NP1s ($z = 1.86$, $SE = 0.30$, $p < .10$), accuracy for actives did not differ across contexts ($z = 0.19$, $SE = 0.21$, $p > .80$). This led to a main effect of construction ($X^2(1, N = 40) = 10.80$, $p < .001$) and an interaction with NP1 status ($X^2(1, N = 40) = 6.45$, $p < .01$).

Next, we turned to role-reversal errors, which suggested that children's difficulties with passives reflect a failure to revise an agent-first bias. Final models included random slopes and intercepts for both subjects and items. Role-reversal errors were more prevalent for passives ($M = 45\%$, $SD = 33\%$) compared to actives ($M = 22\%$, $SD = 18\%$) in sentences with definite NP1s ($z = 2.15$, $SE = 0.23$, $p < .05$), but no reliable construction differences were found with pronoun NP1s ($z = 0.72$, $SE = 0.16$, $p > .40$). Likewise, while role-reversal errors for passives were greater with definite NP1s compared to pronoun NP1s ($z = 2.09$, $SE = 0.26$, $p < .05$), they did not differ for actives ($z = 0.83$, $SE = 0.36$, $p > .40$). This led to a main effect of construction ($X^2(1, N = 40) = 6.41$, $p < .01$) and an interaction with NP1 status ($X^2(1, N = 40) = 10.99$, $p < .001$). Ambiguous actions did not vary with NP1 status or construction (all p's $> .10$).

2.3. Discussion

Experiment 1 found that children's interpretation of constructions varies with NP1 status in English, as it does in Mandarin. Children perform worse with passives compared to actives with definite NP1s, which signal new discourse entities, introduce greater interpretive uncertainty, and increase the agent-first bias. This construction asymmetry disappears with pronoun NP1s, which signal given discourse entities, introduce less uncertainty, and decrease the agent-first bias. These patterns are problematic for the language-specific account, which predicted that semantically rich arguments would improve comprehension

Table 2
In Experiment 1, fixed effects (NP1 status × construction) in logistic mixed-effects models regression model of (a) correct, (b) reverse, and (c) ambiguous actions

	Intercept				NP1 status				Construction				NP1 status × Construction			
	β	*SE*	*z*	*p*	β	*SE*	*z*	*p*	β	*SE*	*z*	*p*	β	*SE*	*z*	*p*
Correct	0.52	0.23	2.25	0.02*	0.25	0.23	1.07	0.28	0.29	0.12	2.28	0.02*	0.31	0.12	2.47	0.01*
Reverse	1.02	0.20	4.96	0.01*	0.17	0.19	0.91	0.36	0.16	0.15	1.09	0.27	0.40	0.14	2.91	0.01*
Ambiguous	2.94	0.35	8.41	0.01*	0.21	0.29	0.72	0.47	0.25	0.18	1.34	0.18	0.25	0.18	1.34	0.18

*Note. $p < .05$ (two tailed)

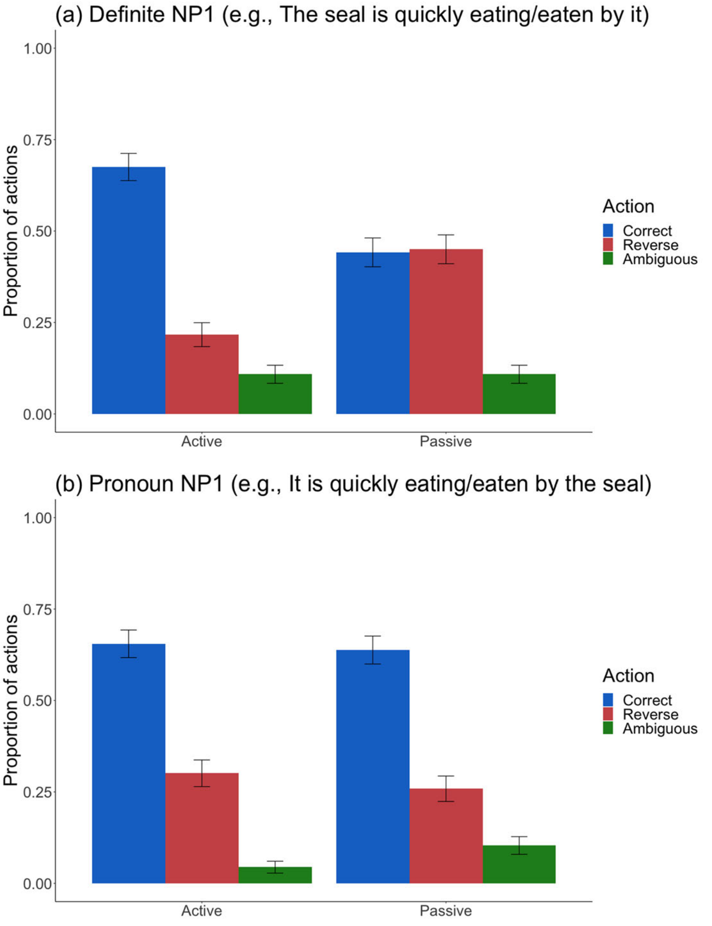

Fig. 1. In Experiment 1, the proportion of actions coded for role assignment in (a) Definite NP1 and (b) Pronoun NP1 conditions. Error bars represent standard error of the mean.

of passives in English. These results, however, are consistent with alternative accounts that explain pronoun-NP1 advantages in terms of expression frequency and/or referential ambiguity.

To distinguish these accounts, Experiment 2 focuses on sentences with pronoun NP1s (e.g., *It is quickly…*) and varies discourse status within the current communicative context, rather than linguistic expressions (e.g., *the seal* vs. *it*). In sentences like (6a), a single NP introduces one prominent referent into the discourse (e.g., a gray seal). After this sentence, pronoun NP1s in (7) naturally refer to this previously mentioned (given) entity (Hartshorne et al., 2015; Song & Fisher, 2005). In contrast, in sentences like (6b), a conjoined NP introduces two prominent referents (i.e., gray seal, white seal). After this sentence, children may be more likely to consider previously unmentioned (new) entities as pronoun-NP1 referents in (7). This follows from Grice's Quantity Maxim (Grice, 1975): If the speaker wished to reference a mentioned entity, she would have used a definite NP to distinguish which one (Arnold, 2010; Givon, 1983). However, since she did not, it is possible that she is referring to another entity in the scene (i.e., exophoric reference, *it* refers to shark or fish).

In Experiment 1, fixed effects (NP1 status × construction) in logistic mixed-effects models regression model of (a) correct, (b) reverse, and (c) ambiguous actions

(6)	a. Single NP:	The gray seal swims
	b. Conjoined NP:	The gray seal and the white seal swim
(7)	a. Active:	It is quickly catching the white seal
	b. Passive:	It is quickly caught by the white seal

By varying discourse status through a prior context, we generate three predictions for relationships between NP1 status and construction. If children recruit an agent-first bias to overcome uncertainty introduced by new discourse entities, agent-first predictions should increase in (7) when conjoined-NP primes promote consideration of previously unmentioned objects. This should lower accuracy for passives compared to actives. Conversely, when single-NP primes encourage reference to a given entity in (7), children may avoid an agent-first bias and interpret roles using verb morphology instead. This should lead to comparable accuracy for actives and passives. If, however, construction asymmetries are mediated by NP1 frequency, then pronoun NP1s should always facilitate retrieval from the lexicon and lead to high accuracy across constructions (similar to pronoun NP1s in Experiment 1). Finally, if the agent-first bias is mediated by referential ambiguity, then construction asymmetries may appear when single-NP primes enable reference restriction, but diminish when conjoined-NP primes block it.

3. Experiment 2

3.1. Methods

3.1.1. Participants

Forty English-speaking children were recruited from private schools in the Washington, DC metro area (18 male, 21 female). The mean age was 5.11 (*SD* 0.9, range = 4.5–7.2). Half the children were randomly assigned to the conjoined-NP condition, and the other half to the single-NP condition. There was no significant difference in age (in months) or gender across the two groups ($p > .80$).

3.1.2. Procedure and materials

The procedure and materials were similar to Experiment 1, but differed in the following ways. First, displays featured four-object sets instead of three. In addition to a likely agent (e.g., shark) and a likely theme (e.g., fish), the mentioned object in critical sentences (e.g., white seal) was paired with a competitor object (e.g., gray seal), which came from the same basic-level category (e.g., seals, people), but varied along salient dimensions (e.g., color, occupation). To ensure that pronouns did not disambiguate referents in critical sentences, within-set objects shared animacy and gender (i.e., *it*, *he*, *she*). Second, prime trials like (6) always appeared as first sentences, and critical trials like (7) as second sentences. Prime sentences were intransitive and varied the number of entities introduced into the discourse. In the single-NP condition, competitor objects performed actions alone (e.g., *The gray seal swims*). In the conjoined-NP condition, competitors were described before mentioned objects, and both performed actions simultaneously (e.g., *The gray seal and the white seal swim*). Critical sentences featured pronoun NP1s and definite NP2s in actives or passives (e.g., *It is quickly catching/caught by the white seal*). Main verbs described actions where competitors could be plausible agents or themes of actions with mentioned objects (e.g., *catching*, *pushing* but not *eating*). This ensured that mentioned objects were potential referents of pronoun NP1s in critical sentences. Twelve prime and critical trials were presented with eight additional filler trials. See Appendix B for a full list of items.

3.1.3. Coding

Similar to Experiment 1, only actions for critical sentences were coded. Approximately 2.1% of trials were excluded because of experimenter error or because no action was produced. The remaining actions were coded along two dimensions. First, to examine whether prime sentences influenced which entities were considered as potential referents for pronoun NP1s, research assistants coded which objects were selected in critical sentences besides the mentioned entity (e.g., white seal). Actions involving likely agents or likely themes (e.g., shark or fish) were coded as *new referents*. Those involving competitor objects (e.g., gray seal) were coded as *given referents*. Approximately 6.0% of trials were excluded from these analyses because no additional object was selected (i.e., actions only involved the mentioned object) or because new and given referents were both selected.

Second, to assess the accuracy of role assignment, actions were recategorized based on relationship to the construction. *Correct actions* depicted accurate role assignments between the mentioned object and another plausible object. For active trials, these involved making the likely agent or competitor object do something to the mentioned object (e.g., a shark or gray seal catching the white seal). For passive trials, they involved the mentioned object doing something to the likely theme or competitor object (e.g., making the white seal catch a fish or gray seal). *Reverse actions* depicted incorrect role assignments between the mentioned object and another plausible object. For actives, these involved the mentioned object doing something to the likely themes or competitor object. For passives, they involved the likely agent or competitor object doing something to the mentioned object. *Ambiguous actions* were ones where only the mentioned object was selected or all four objects were selected. To ensure reliability in coding, a second research assistant confirmed the original coding on 96.5% of trials.

3.2. Results

Fig. 2 illustrates that pronoun NP1s were often interpreted as referring to competitor objects. This is unsurprising since they were always mentioned in prime sentences. We analyzed responses using logistic mixed-effects models, with prime sentence (conjoined-NP vs. single-NP) and construction (active vs. passive) modeled as fixed-effects variables. Final models included random intercepts for subjects but omitted them for items. Model comparison revealed no significant difference in the variance explained by models that included items as random effects (p's $> .50$). As predicted, children produced more actions involving new referents (i.e., likely agents, likely themes) following conjoined-NP primes ($M = 33\%$, $SD = 22\%$) compared to single-NP primes ($M = 4\%$, $SD = 8\%$; $X^2(1, N = 40) = 33.98$, $p < .001$). There was no additional effect or interaction with construction (all p's $> .30$). These results confirm that conjoined-NP primes increase the degree to which new entities are considered as potential referents of pronoun NP1s.

Next, we examined the accuracy of role assignment in critical sentences. Final models included random intercepts for subjects and random slopes and intercepts for items. Model comparison revealed no significant difference in the variance explained by maximal models (p's $> .40$). Fig. 3 illustrates that prime sentences also impacted the accuracy of role assignment in critical sentences. Children were less accurate with passives ($M = 54\%$, $SD = 31\%$) compared to actives ($M = 91\%$, $SD = 10\%$) when preceded by conjoined-NP primes ($z = 4.35$, $SE = 0.57$, $p < .001$), but this difference diminished after single-NP primes ($z = 1.51$, $SE = 1.51$, $p > .10$). Likewise, while passives were less accurate after conjoined-NP primes compared to single-NP primes ($z = 2.79$, $SE = 0.70$, $p < .01$), accuracy for actives did not reliably vary across contexts ($z = 0.18$, $SE = 0.49$, $p > .80$). This led to main effects of prime sentence ($X^2(1, N = 40) = 9.29$, $p < .01$) and construction ($X^2(1, N = 40) = 37.68$, $p < .001$) and an interaction between the two ($X^2(1, N = 40) = 7.12$, $p < .01$).

Finally, we turned to role-reversal errors and ambiguous actions. Final models included random intercepts for subjects and random slopes and intercepts for items, and model comparison revealed no significant difference in the variance explained by models that included items as random effects (p's $> .20$). Role-reversal errors were greater after conjoined-NP primes compared to single-NP primes ($X^2(1, N = 40) = 12.11$, $p < .001$) and for passives compared to actives ($X^2(1, N = 40) = 15.18$, $p < .001$). Unlike Experiment 1, there was no interaction between prime sentence and construction ($X^2(3, N = 40) = 0.10$, $p > .70$). Instead, ambiguous actions now revealed a prime by construction interaction ($X^2(3, N = 40) = 7.29$, $p < .01$). They were more prevalent for passives ($M = 11\%$, $SD = 17\%$) compared to actives ($M = 3\%$, $SD = 6\%$) after conjoined-NP primes ($z = 2.36$, $SE = 0.74$, $p < .05$), but there was no reliable difference after single-NP primes ($z = 1.11$, SE $= 0.63$, $p > .20$). Similarly, comparisons within constructions revealed that ambiguous actions were greater after conjoined-NP primes compared to single-NP primes for passives ($z = 2.01$, SE $= 0.79$, $p < .05$), but did not differ reliably for actives ($z = 0.94$, $SE = 0.83$, $p > .30$).

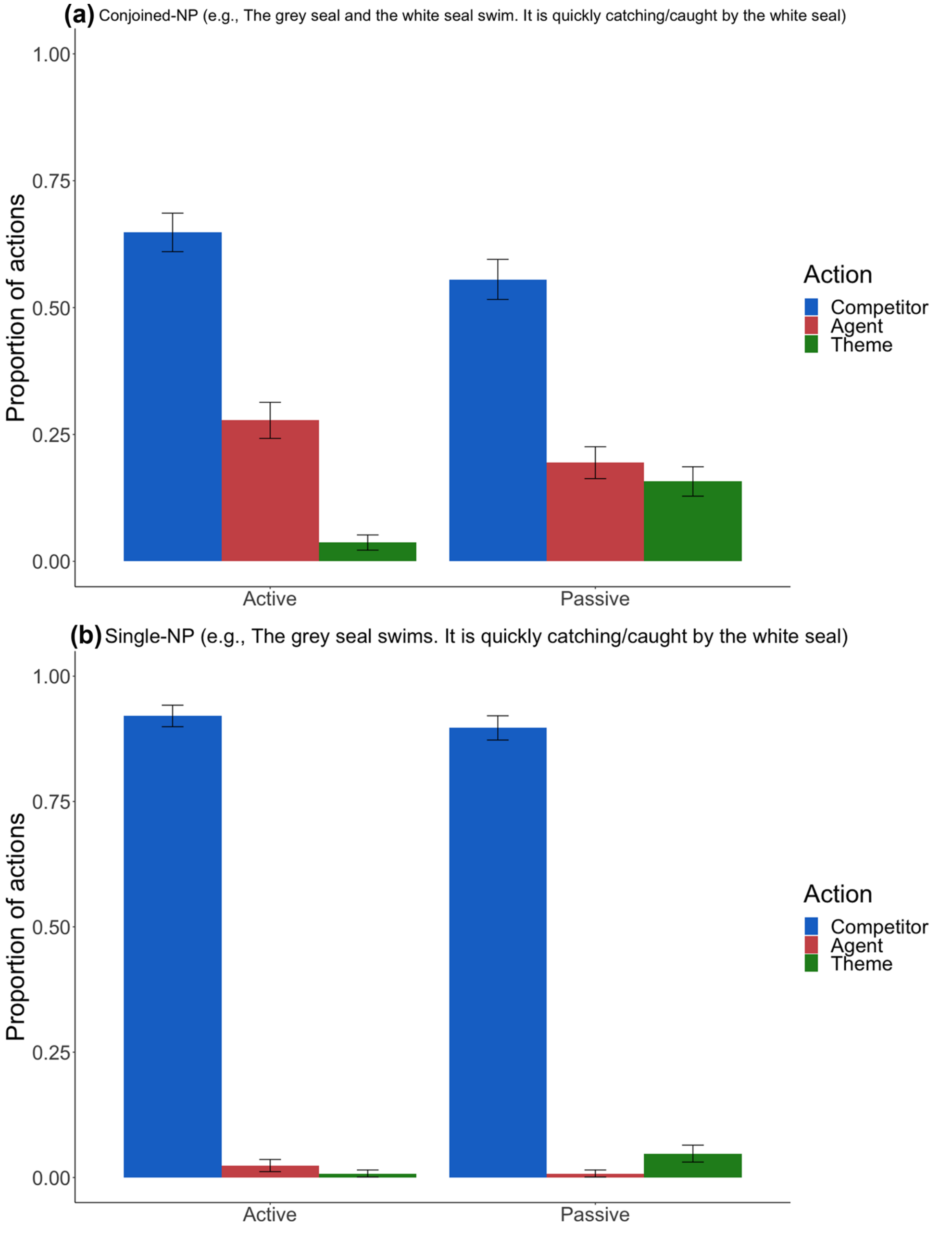

Fig. 2. In Experiment 2, the proportion of actions coded for pronoun-NP1 identity in (a) Conjoined-NP and (b) Single-NP prime conditions. Error bars represent standard error of the mean.

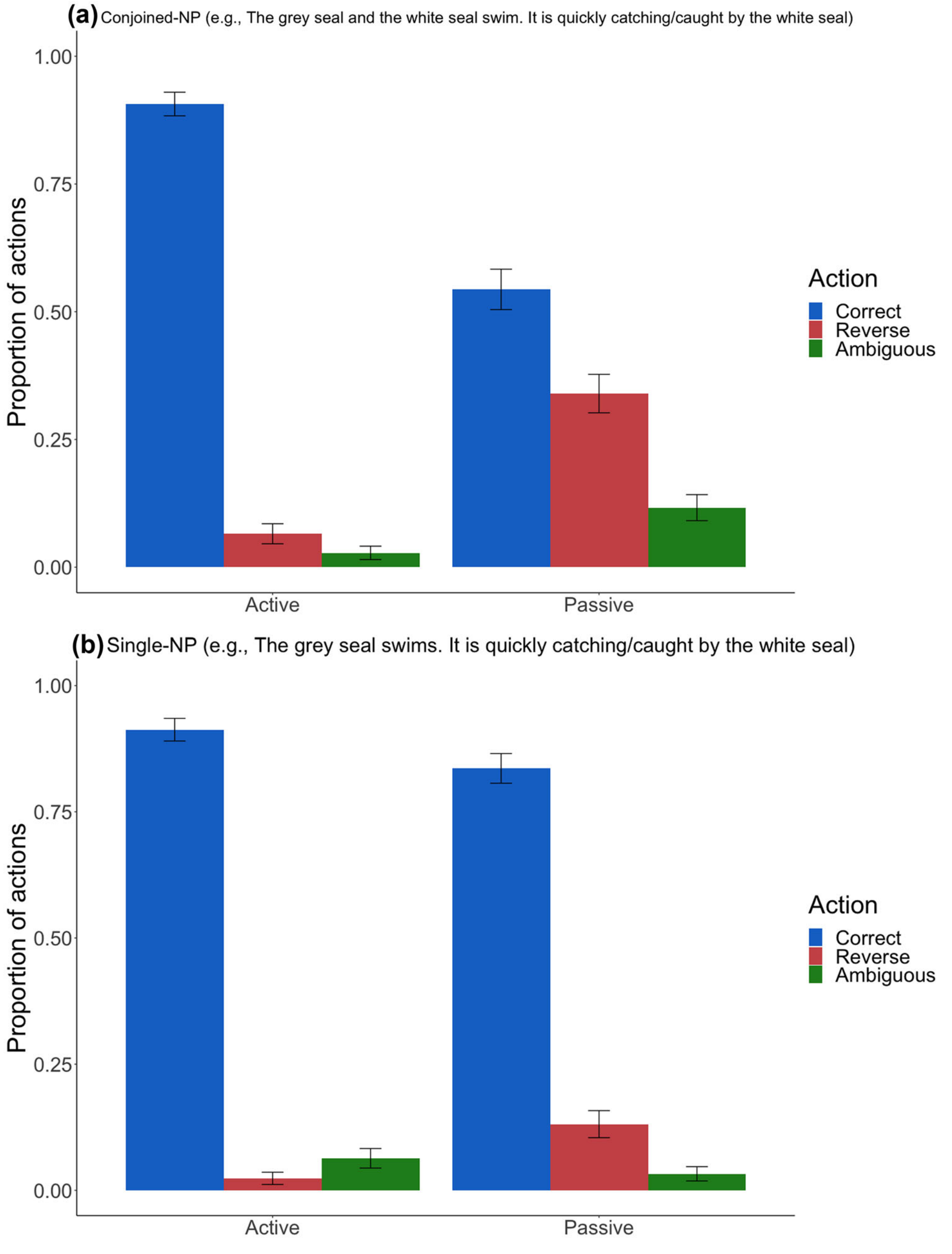

Fig. 3. In Experiment 2, the proportion of actions coded for role assignment in (a) Conjoined-NP and (b) Single-NP prime conditions. Error bars represent standard error of the mean.

Table 3
In Experiment 2, fixed effects (prime sentence × construction) in logistic mixed-effects models regression model of actions with unmentioned objects

	β	*SE*	*z*	*p*
Intercept	2.01	0.21	9.44	0.01^{*}
Prime	1.26	0.20	6.15	0.01^{*}
Construction	0.19	0.18	1.09	0.28
Prime x Construction	0.10	0.18	0.58	0.56

Note. $^{*}p < .05$ (two tailed)

Table 4
In Experiment 2, fixed effects (prime sentence × construction) in logistic mixed-effects models regression model of (a) correct, (b) reverse, and (c) ambiguous actions

	Intercept				Prime				Construction				Prime× Construction			
	β	*SE*	*z*	*p*	β	*SE*	*z*	*p*	β	*SE*	*z*	*p*	β	*SE*	*z*	*p*
Correct	2.12	0.35	6.04	0.01*	0.45	0.20	2.18	0.03*	1.09	0.32	3.38	0.01*	0.39	0.15	2.50	0.01*
Reverse	2.62	0.39	6.76	0.01*	0.64	0.23	2.77	0.01*	1.14	0.32	3.62	0.01*	0.06	0.20	0.32	0.75
Ambiguous	3.85	0.62	6.23	0.01*	0.16	0.28	0.56	0.57	0.29	0.53	0.54	0.59	0.62	0.24	2.54	0.01*

*Note. $p < .05$ (two tailed)

3.3. Discussion

Experiment 2 found that communicative contexts that promote reference to new entities also increase the agent-first bias. This, in turn, diminished comprehension of passives relative to actives. In contrast, children interpreted passives as accurately as actives when pronoun NP1s referred to given entities, which reduced uncertainty and the agent-first bias. This is striking since critical sentences were identical across prime conditions, thus variation in comprehension cannot be explained by lexical properties that mediate retrieval from the lexicon or integration into sentence structures (e.g., frequency, semantic richness). These results are also inconsistent with the referential-ambiguity account. If failure to restrict reference blocks an agent-first bias, then passives should have been more accurate when NP1s were ambiguous after conjoined-NP primes (i.e., *it* refers to competitors, likely agents, likely themes) compared to unambiguous after single-NP primes (i.e., *it* refers to competitors on >90% of trials). Evidence of the opposite suggests that the agent-first bias is triggered by the uncertainty of discourse status, rather than the certainty of reference restriction.

One potential concern is that the conjoined-NP primes created an infelicitous context, and children's decreased accuracy reflected general confusion about how to interpret sentences under these conditions. To address this, we presented our materials to 48 adults and assessed the extent to which they interpreted pronouns and sentences in the expected manner. We found that similar to children, adults were more likely to interpret pronouns as referring to new objects after conjoined-NP primes ($M = 52\%$, $SD = 26\%$) compared to single-NP primes ($M = 8\%$, $SD = 17\%$; $X^2(1, N = 48) = 38.07$, $p < .001$). This confirms that conjoined-NP primes increase the degree to which new entities are considered as referents of pronoun NP1s. Adults were also highly accurate at assigning roles in critical sentences (>95% across conditions), and were more accurate after single-NP primes compared to conjoined-NP primes ($X^2(1, N = 48) = 8.04$, $p < .01$). There was no effect or interaction with construction (all p's $> .20$). While these data do not rule out the possibility that children's interpretation was impacted by infelicitous aspects of the conjoined-NP primes, they demonstrate that the current task and materials are sensitive to linguistic knowledge for coreferencing (to interpret pronouns) and role assignment (to interpret actives and passives), and developmental intuitions about the former mirror adult ones.

In Experiment 3, we returned to definite NP1s (e.g., *The seal*), and examined a context that should reduce the agent-first bias for these expressions. We reasoned that if the agent-first bias reflects a statistical tendency for definite NPs to refer to newer discourse entities compared to pronouns in Experiment 1, then it may be weakened by even newer entities. By definition, novel words (e.g., *The blicket*) offer strong cues to new entities (e.g., mutual-exclusivity effects in Bion, Borovsky, & Fernald, 2013; Markman & Wachtel, 1988 among others). To examine their interpretation, we divided trials into two parts. During the familiarization phase, children saw a familiar object (e.g., seal) interacting with two unfamiliar objects. For example, a large monster-like predator (i.e., likely agent) chases a seal, and a seal chases a small, wimpy prey (i.e., likely theme). During the test phase, children heard actives and passives containing a novel word like in (8) and (9), and chose a corresponding object (e.g., *Click on the blicket!*). Object selection reveals how children assigned roles to familiar nouns (e.g., *the seal*) and inferred referents for novel words. For Novel NP1s in (8), familiar nouns are themes in actives, implying that novel words are likely agents. Conversely, familiar nouns are agents in passives, implying that novel words are likely themes. For Known NP1s in (9), interpretation of novel words should generate a preference for likely themes in actives and likely agents in passives.

In Experiment 2, fixed effects (prime sentence × construction) in logistic mixed-effects models regression model of (a) correct, (b) reverse, and (c) ambiguous actions

(8)	Novel NP1	a. Active:	The blicket will be quickly eating the seal
		b. Passive:	The blicket will be quickly eaten by the seal
(9)	Known NP1	a. Active:	The seal will be quickly eating the blicket
		b. Passive:	The seal will be quickly eaten by the blicket

If the uncertainty of new discourse entities mediates the agent-first bias, then children should be less accurate with passives compared to actives following novel NP1s in (8). In contrast, known NP1s in (9) may decrease relative uncertainty and delay role assignment until the late-arriving verb. This would lead to high accuracy for actives and passives alike. If comprehension is instead tied to the ease of retrieving NP1 expressions from the lexicon, then known NP1s should lead to more difficulty with passives than actives, much like definite NP1s did in Experiment 1. A similar pattern may emerge with novel NP1s, which have a prior frequency of zero. Alternatively, if referential ambiguity blocks a default tendency to adopt the agent-first bias, then novel NP1s may improve comprehension of passives by delaying reference restriction until the main verb (e.g., the role of *The blicket* is undetermined until *eating* or *eaten by*). In contrast, known NP1s can be immediately associated with familiar referents in displays, triggering an agent-first bias and lowering accuracy for passives compared to actives.

4. Experiment 3

4.1. Methods

4.1.1. Participants

Forty English-speaking children were recruited from private schools in the Washington, DC metro area (21 male, 19 female). The mean age was 5.4 ($SD = 0.3$, range = 5.0–5.11). Half of the children were randomly assigned to the novel NP1 condition, and the other half to the known NP1 condition. There was no significant difference in age (in months) or gender across the two groups ($p > .50$).

4.1.2. Procedure and materials

Unlike Experiments 1 and 2, this experiment used computer displays rather than live objects. Children were told that they were going to play a game where they picked familiar and unfamiliar objects. Familiar objects were people and animals that are well known to children (e.g., a seal). Unfamiliar objects were novel objects that were introduced during the study. Each trial featured two parts. During the familiarization phase, the familiar object was presented in the middle of the screen and linguistically mentioned (e.g., *Look at the seal!*). This increased its status as a given entity relative to unfamiliar objects (Arnold, 2010). This event was followed by unlabeled events where familiar and unfamiliar objects interacted with each other. Likely agents acted on familiar objects (e.g., a large, menacing novel creature chasing the seal), and familiar objects acted on unfamiliar, likely themes (e.g., the seal chasing a small, puny novel creature).

During the test phase, children saw unfamiliar objects on either side of familiar objects and heard sentences describing an up-coming event like (8) and (9). All sentences included novel and known words but varied their order of occurrence varied with NP1 status. Novel NP1s (e.g., *The blicket is…*) featured novel words in subject position while known NP1s (i.e., *The seal is…*) featured novel words in object position. Construction type contrasted active versus passive sentences. Twelve critical trials were randomized with six filler trials, which embedded familiar words in active sentences (e.g., *The sheep will be slowly eating the grass. Click on the grass.*). See Appendix C for a complete list of items.

4.1.3. Coding

Children's selection of objects for critical trials was coded based on the condition. *Correct actions* involved selecting an unfamiliar object based on correct role assignment of the familiar noun. For novel NP1/active and known NP1/passive trials, this referred to the likely agent. For novel NP1/passive and known NP1/active trials, this referred to the likely theme. *Reverse actions* involved selecting an unfamiliar object based on incorrect role assignment of the familiar noun. For novel NP1/active and known NP1/passive trials, this referred to the likely theme. For novel NP1/ passive and known NP1/active trials, this referred to the likely agent. *Ambiguous actions* involved selecting the familiar object.

4.2. Results

We first examine the accuracy of children's interpretation. Children's responses were analyzed using logistic mixed-effects models with NP1 status (novel vs. known NP1) and construction (active vs. passive) as fixed effects. Final models included random slopes and intercepts for subjects but only intercepts for items. Model comparison revealed no significant difference in the variance explained by maximal models (p's $> .30$). Fig. 4 illustrates that passives ($M = 39\%$, $SD = 34\%$) were less accurate than actives ($M = 87\%$, $SD = 17\%$) for novel NP1s ($z = 3.71$, $SE = 0.89$, $p < .001$), but reliable differences across constructions were not found for known NP1s ($z = 1.15$, $SE = 0.71$, $p > .20$). Similarly, while passives were less accurate with novel NP1s compared to known NP1s ($z = 3.26$, $SE = 0.66$, $p < .01$), this pattern reversed for actives ($z = 2.62$, $SE = 0.81$, $p < .01$). This led to a main effect of construction ($X^2(1, N = 40) = 19.23$, $p < .001$) and an interaction with NP1 status ($X^2(1, N = 40) = 78.91$, $p < .001$).

Next, we turned to role-reversal errors, which revealed parallel patterns. Final models included random slopes and intercepts for subjects but only intercepts for items. Model

Table 5
In Experiment 3, fixed effects (NP1 status × construction) in logistic mixed-effects models regression model of (A) correct and (B) reverse actions

	Intercept				NP1 status				Construction				NP1 status× Construction			
	β	*SE*	*z*	*p*	β	*SE*	*z*	*p*	β	*SE*	*z*	*p*	β	*SE*	*z*	*p*
Correct	0.97	0.23	4.13	0.01^{*}	0.14	0.21	0.69	0.49	0.48	0.21	2.27	0.02^{*}	0.95	0.21	4.52	0.01^{*}
Reverse	1.13	0.24	4.78	0.01^{*}	0.01	0.20	0.02	0.98	0.55	0.23	2.46	0.01^{*}	1.03	0.22	4.63	0.01^{*}

**Note*. $p < .05$ (two tailed)

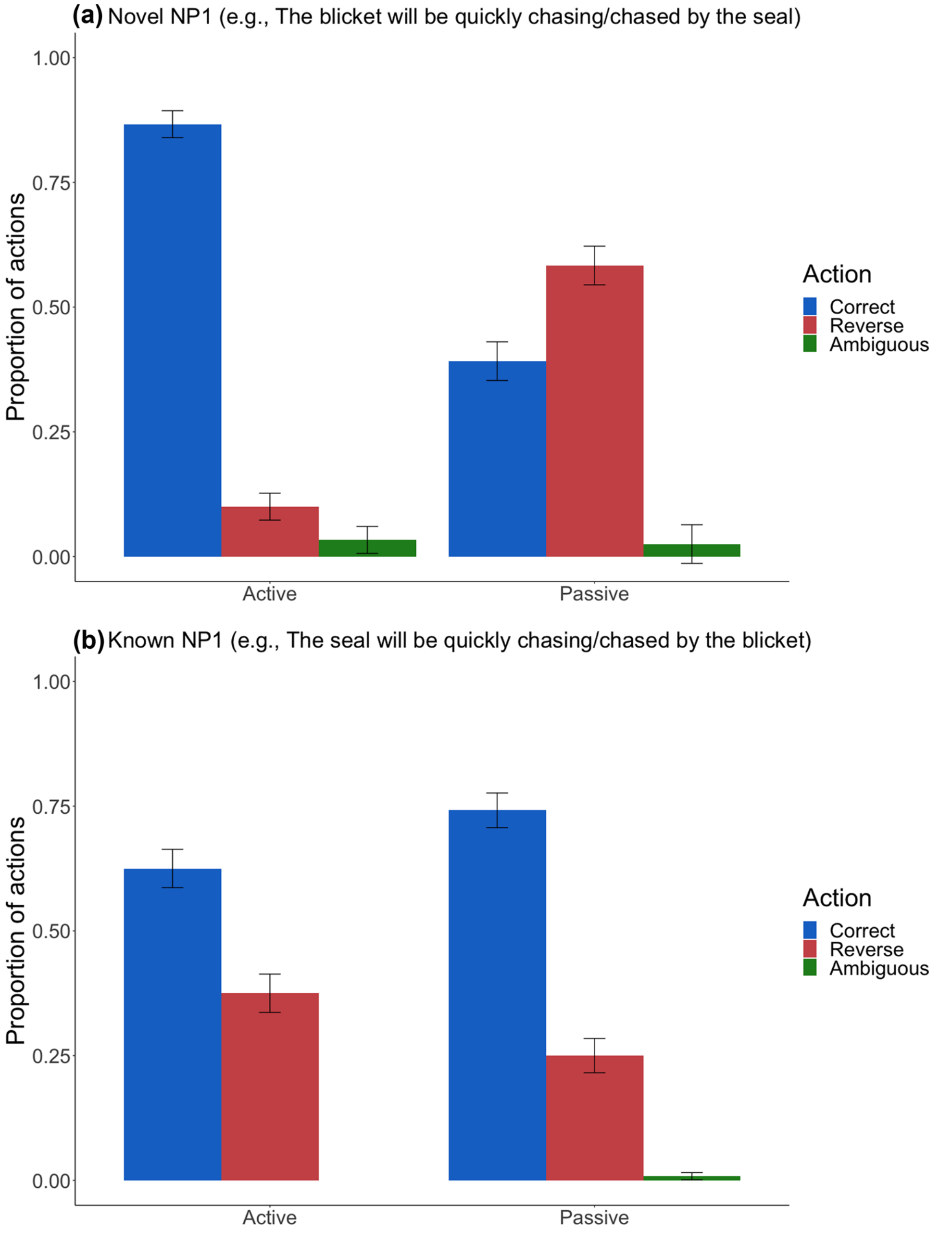

Fig. 4. In Experiment 3, the proportion of actions coded for role assignment in (a) Novel NP1 and (b) Known NP1 conditions. Error bars represent standard error of the mean.

Table 6
In Experiments 1–3, fixed effects (NP1 status × experiment) in logistic mixed-effects models regression model of correct actions for (a) active and (b) passive constructions

	Active				Passive			
	β	SE	z	p	β	SE	z	p
Intercept	2.57	0.29	8.86	0.01*	1.21	0.31	3.87	0.01*
Experiment 1	1.77	0.35	5.12	0.01*	1.00	0.43	2.34	0.02*
Experiment 3	1.21	0.35	3.40	0.01*	0.76	0.43	1.77	0.08
NP1 status	0.04	0.27	0.16	0.87	0.99	0.31	3.16	0.01*
Experiment 1: NP1 status	0.08	0.34	0.25	0.80	0.43	0.42	1.02	0.30
Experiment 3: NP1 status	0.80	0.35	2.27	0.02*	0.05	0.43	0.12	0.90

*Note. $p < .05$ (two tailed)

comparison revealed no significant difference in the variance explained by maximal models (p's $> .20$). They were more frequent for passives ($M = 58\%$, $SD = 33\%$) compared to actives ($M = 10\%$, $SD = 16\%$) for novel NP1s ($z = 3.61$, $SE = 0.99$, $p < .001$), but there were no reliable differences across constructions for known NP1s ($z = 1.22$, $SE = 0.69$, $p > .20$). Similarly, while reversal-errors were more frequent with novel compared to known NP1s for passives ($z = 3.19$, $SE = 0.63$, $p < .01$), this pattern reversed for actives ($z = 3.15$, $SE = 0.75$, $p < .01$). This led to a main effect of construction ($X^2(1, N = 40) = 20.98$, $p < .001$) and an interaction with NP1 status ($X^2(1, N = 40) = 60.79$, $p < .001$). Ambiguous actions were not analyzed since they occurred on only 1.7% of trials.

4.2.1. Comparing the agent-first bias across linguistic contexts

To understand how the agent-first bias varies across linguistic contexts, we conducted two follow-up analyses across experiments. First, the presence of null construction effects in the given condition raised questions about whether decreased uncertainty weakened the agent-first bias or whether the current study simply lacked power to detect sizable effects. To address this, we turned to Bayesian inference, which has gained ground as an alternative to null-hypothesis significance testing (Jarosz & Wiley, 2014; Lee & Wagenmakers, 2014; van Doorn, Ly, Marsman, & Wagenmakers, 2020). Unlike traditional p-values, this approach calculates the relative odds of two hypotheses about the current data by estimating beliefs about the hypotheses before the data (priors) and updates to beliefs after the data (likelihoods). We used the BayesFactor package 0.9.2 in R (Morey & Rouder, 2011), and separated correct responses by NP1 status (given vs. new) and estimated the inverse Bayes factors (null/alternative) for models that included and excluded construction.

Across NP1 status (given vs. new), comparisons of the inverse Bayes Factors (BF10s) reveal how the accuracy of interpretation is affected by construction (actives vs. passives) and the extent to which this holds across experiments. When NP1s referred to new entities, the data were more likely to occur under models that included construction effects, compared to models without them. Specifically, the alternative hypothesis was 2.7 times more likely in Experiment 1 (positive evidence), 587.1 times more likely in Experiment 2 (very strong evidence), and 7063.5 times more likely in Experiment 3 (very strong evidence). When NP1s referred to given entities, the data were far less likely to occur under models that included construction effects, compared to models without them. Specifically, the alternative hypothesis was 0.32 times more likely to occur under a model with construction effects in Experiment 1 (weak evidence), 0.74 times more likely in Experiment 2 (weak evidence), and 0.58 times more likely in Experiment 3 (weak evidence). Together, this suggests that construction effects are qualitatively different across contexts that elicit an agent-first bias compared to those that do not.[1]

Next, to test the possibility that children recruit the agent-first bias to tackle tasks uncertainty generally, rather than discourse status specifically, we separated correct responses by construction and created models with NP1 status (given vs. new) and Experiment as fixed effects. Fig. 5 illustrates that for passives, prime sentences in Experiment 2 increased the overall accuracy of role assignment relative to Experiment 1 ($z = 2.34$, $SE = 0.43$, $p < .05$) and Experiment 3 ($z = 1.77$, $SE = 0.43$, $p < .10$). However, correct responses were always greater when NP1s referred to given compared to new entities ($X^2(1, N = 40) = 23.26$, $p < .001$). There was no interaction with experiment (p's $> .40$). This suggests that difficulties with passives arise when communicative contexts incur an agent-first bias, and this was independent of task-specific properties. For actives, accuracy was again greater for Experiment 2 compared to Experiment 1 ($z = 5.12$, $SE = 0.35$, $p < .001$) and Experiment 3 ($z = 3.40$, $SE = 0.35$, $p < .001$). Accuracy was greater when NP1s referred to new entities compared to given ones in Experiment 3, and this discourse effect was greater compared to Experiment 1 ($z = 2.39$, SE $= 0.30$, $p < .05$) and Experiment 2 ($z = 2.27$, SE $= 0.35$, $p < .05$). This led to an interaction between NP1 and construction ($X^2(1, N = 40) = 7.49$, $p < .05$). Since new entities correspond to novel NP1s in Experiment 3 (e.g., *The blickets*), these findings raise the possibility that the agent-first bias may lead children to adopt default active interpretations when existing linguistic knowledge is absent.

4.3. Discussion

Experiment 3 provides converging evidence that children's recruitment of the agent-first bias is linked to the discourse status of NP1s, rather than specific expressions. When novel NP1s referred to unfamiliar entities, accuracy with passives was lower than actives. However, when known NP1s referred to familiar entities, accuracy was similar across constructions. Notably, identical NP1s (e.g., *the seal*) generated variable role assignments in relation to NP2 expressions in Experiments 1 and 3. When definite NP1s signaled newer entities relative to pronoun NP2s (e.g., *it*), children adopted an agent-first bias. When known NP1s signaled given entities relative to novel NP2s (e.g., *the blicket*), they withheld this bias. These findings are inconsistent with NP1 frequency, which fails to explain why interpreting known NP1s is easier when they occur with novel NP1s compared to pronoun NP1s. Similarly, if the absence of reference restriction blocks the agent-first bias, then novel NP1s are initially consistent with either unfamiliar object and this should have facilitated processing of passives. Instead, decreased accuracy in this context suggests that children readily assign agent roles to ambiguous expressions.

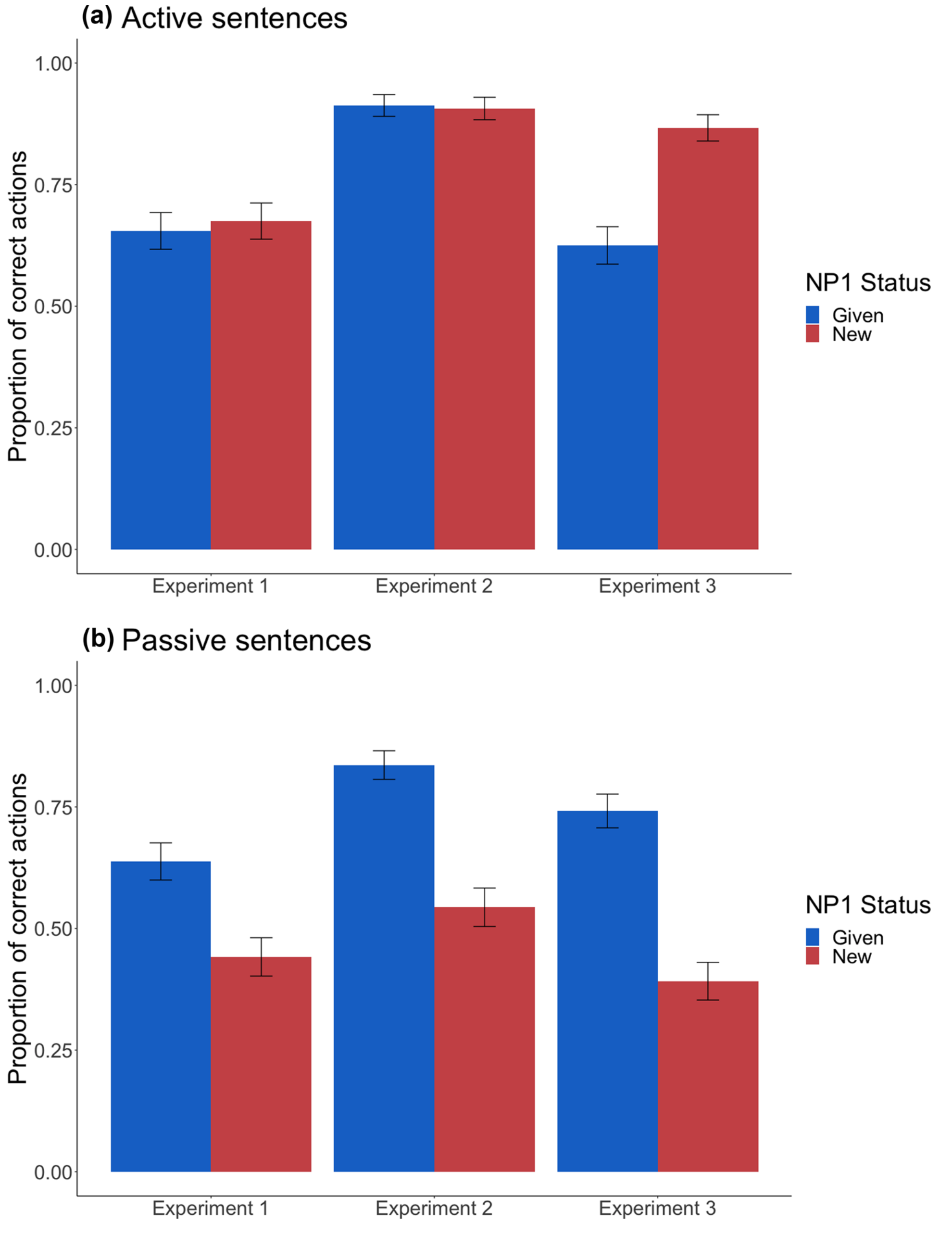

Fig. 5. Across Experiments, the proportion of correct actions based on the discourse status of NP1 in (a) active (b) passive conditions. Error bars represent standard error of the mean.

5. General Discussion

This study investigated the development of syntactic parsing by examining why children sometimes recruit an agent-first bias and sometimes not. In three experiments, we found that the agent-first bias is related to discourse status, implied through linguistic cues such as the statistical properties of expressions (e.g., definite NPs vs. pronouns), relations to the preceding sentence (e.g., exophoric vs. anaphoric), introduction of novel entities (e.g., novel vs. known NPs). When linguistic contexts signal that NP1s are new entities, this increases interpretive uncertainty and the agent-first bias, which improves comprehension of actives, but hinders passives. In contrast, when NP1s are given entities, this decreases uncertainty, weakens the bias, and leads to accurate comprehension across constructions. These findings suggest that children draw on parsing cues in a heuristic-like manner, exploiting broad-scale regularities (e.g., an agent-first bias) when uncertainty is high but switching to finer-grained properties (e.g., verb morphology) when uncertainty is low.

At a computational level, this strategy solves basic challenges that children face about how to parse their input, given varying sentence properties and linguistic knowledge. If the child is too conservative and eschews predictions altogether, they may fail to interpret roles within fast-moving sentences. However, if the child is too liberal and generates predictions based solely on cue order rather than reliability, they may face widespread kindergarten-path effects. Thus, like adults (Ferreira et al., 2002; Gibson et al., 2013; Levy et al., 2009), children may optimize how they calculate sentence meanings across communicative contexts. This is consistent with recent evidence of parsing variability in children. For example, Ovans, Novick, and Huang (2020) finds that cognitive-control engagement leads children to ignore early-arriving cues and assign roles using reliable cues instead. Likewise, recent work demonstrates that the agent-first bias plays an important role in sentence interpretation. Unlike typically developing peers, children with developmental language disorders are less likely to recruit the agent-first bias, and this limits comprehension of active sentences when semantic/plausibility cues are absent (Montgomery, Gillam, Evans, & Sergeev, 2017; Oppenheimer et al., 2020).

A potential way to implement variable parsing strategies is as a tradeoff between the predictability of discourse referents (where new referents are less predictable than given ones) and reliance on language-general parsing statistics (where the agent-first bias is a prior computed over all sentences while verb biases are likelihoods over specific ones). During comprehension, the child may track entities in the discourse based on the visual scene and previous mentioned. It may be possible to quantify these computations using metrics that capture uncertainty such as surprisal (Hale, 2001; Levy, 2008). While surprisal is generally used to measure the likelihood that a syntactic parse is ruled out by prior words (e.g., Ovans, Huang, & Feldman, 2020), it may be extended to quantify the negative log probability of referents given prior context. Discourse surprisal may be one input into a parsing system for up- and down-weighting parsing cues given the current context (e.g., noisy-channel parsing), and this may explain how children's parsing strategies are regulated by in-the-moment demands from higher-level communication (e.g., predictability of referents). Similar appeals have been made about the relations between discourse topics and word learning (Roy, Frank, DeCamp, Miller,

& Roy, 2015), word prediction (Bhattasali & Resnik, 2020), and phonetic learning (Frank, Feldman, & Goldwater, 2014).

In the remainder of this Discussion, we will focus on three additional issues related to the current findings. First, we will consider the extent to which our findings reflect the pragmatics of passives, rather than relations to incremental role assignment. Second, we will reconcile our findings with an extensive literature indicating that children ignore discourse context during syntactic parsing (Snedeker & Trueswell, 2004; Trueswell et al., 1999; Kidd & Bavin, 2005; Choi & Trueswell, 2010; Weighall, 2008; Woodard et al., 2016), and consider more broadly how discourse cues and sentence structures may interact within parental input during acquisition. Third, we will discuss the implications of our findings for understanding the dynamics of moment-to-moment comprehension versus year-to-year development, and evaluate the extent to which parsing heuristics may mediate relationships between the two.

5.1. *Can pragmatic infelicity explain the current findings?*

We have argued that comprehending passives is difficult when it requires listeners to revise agent-first predictions, which vary with the uncertainty of different discourse entities. However, it is possible that the patterns we observed reflect instead relations between the pragmatics of passives, input statistics, and predicting likely constructions. During communication, speakers use passives to highlight topical themes in subject position (Johnson-Laird, 1968; Williams, 1977), and pronouns to refer to prominent entities in the discourse (Arnold, 2010; Givon, 1983). Thus, children may expect passives to occur more often with pronoun NP1s compared to definite NP1s in input statistics, and have difficulty interpreting passives when NP1s refer to new entities in the current study.

We reject this account for two reasons. First, the infrequency of passives makes relations to NP1 expressions unlikely to be unhelpful for predicting likely roles. Following Huang and colleagues (2017), we analyzed 2,467 spoken utterances in the British National Corpus that resembled the structure of the current stimuli (i.e., NP1 *be* VP-*ing* or VP-*ed by* NP2) (Aston & Burnard, 1998). Consistent with pragmatic usage, passives occur more often with pronoun NP1s ($n = 137$) compared to definite NP1s ($n = 16$). However, this is swamped by asymmetries between actives ($n = 2314$) over passives ($n = 153$) and pronoun NP1s ($n = 1940$) over definite NP1s ($n = 527$). To predict constructions based on these statistics, children can compute the posterior probability of passives given pronoun NP1s (i.e., p(passives|pronoun NP1s) as the product of the likelihood of pronoun NP1s given passives (i.e., p(pronoun NP1s| passives) = 89.5%) and the prior for passives (i.e., p(passives) = 6.2%). They can compare this to the probability of actives given pronoun NP1s (i.e., p(actives|pronoun NP1s), computed as the likelihood of pronoun NP1s given actives (i.e., p(pronoun NP1s|actives) = 77.9%) times the prior for actives (i.e., p(actives) = 93.8%). Importantly, given the vast discrepancy in frequency, actives (73.1%) are far more likely than passives (5.5%), even in pronoun-NP1 contexts. This remains true if passives always appeared with pronoun NP1s and never with definite NP1s. Similar dynamics are present in child-directed speech as well, where passives account for only 0.02% to 3.6% of parental utterances (Gordon & Chafetz, 1990; Laakso

& Smith, 2007). Thus, data sparsity creates challenges for inferring passives based on NP1 distributions alone.

Second, appeals to pragmatic felicity do not predict other facets of our data. It does not explain why children were as accurate with passives as actives following pronoun NP1s in Experiment 1, despite the unpredictability of passives. Likewise, if children assumed that passives are infelicitous with definite NP1s, it is unclear why comprehension improves when known NPs were contrasted with novel NPs in Experiment 3 (e.g., *The seal* vs. *The blicket*). Finally, the pragmatics of passives does not explain why actives are more accurate after novel NP1s compared to known NP1s in Experiment 3. This advantage cannot reflect a visual preference to select more agentive object alone, since bigger referents corresponded to given (pronoun) NP1s in Experiment 1 and did not yield a similar advantage. Instead, these data suggest that novel words may be particularly susceptible to the agent-first bias (e.g., Dautriche et al., 2014), perhaps because they incur extreme interpretive uncertainty. When they appear in sentences, children may maximize their ability to bootstrap likely meanings by relying on canonical parsing cues.

5.2. *Children's sensitivity and insensitivity to discourse context*

To the best of our knowledge, our findings are the first to demonstrate that children simultaneously interpret relations across sentences (e.g., discourse continuity) and within sentences (e.g., role assignment), such that their commitments to the former can influence parsing cues for the latter. This account contrasts with the characterization of developmental parsing as a modular, within-sentence process (Choi & Trueswell, 2010; Kidd & Bavin, 2005; Snedeker & Trueswell, 2004; Trueswell et al., 1999; Weighall, 2008; Woodard et al., 2016). Evidence for this has largely come from a single test case, namely children's failure to infer that contrasting objects in the discourse imply modification (i.e., ḥ for the ambiguous PP, and use a large fan as an instrument on >80% of trials (Snedeker & Trueswell, 2004). When sentences appear with two-referent displays (i.e., pig holding a fan vs. pig holding a fork), adults switch to NP-attachment and infer modifiers to distinguish between referents. On 33% of trials, they use their own hands to tickle the mentioned referent (i.e., pig holding a fan). Children, on the other hand, are unaffected by discourse context, and consistently perform actions based on verb bias alone.

However, it has been noted that correlations between referent number in scenes and post-nominal modification are fairly weak in communication (Huang & Snedeker, 2013; Snedeker & Trueswell, 2004). While adults produce modifiers to disambiguate multiple referents (e.g., *Pick up the square on the left*), they also utter bare NPs (e.g., *Pick up the square*) when referents can be inferred via task goals (Brown-Schmidt & Tanenhaus, 2008). This variability may explain why children are reluctant to adopt a modifier interpretation for VP-biased verbs in two-referent scenes, and why even adults maintain an instrument interpretation on 67% of trials (Snedeker & Trueswell, 2004). Outside of syntactic parsing, there is growing evidence that children exploit linguistic and extra-linguistic cues to discourse continuity to predict sentence meanings. Speakers' production of adjectival modifiers like *big* and *tall* robustly correlates with contrasting referents in a scene (Brown-Schmidt & Tanenhaus, 2006; Ferreira, Slevc, & Rogers, 2005). Consequently, adults and 5-year-olds alike are faster to interpret *Pick up the tall glass* in a two-referent context (i.e., tall glass vs. short glass) compared to a one-referent context (i.e., tall glass vs. short can) (Huang & Snedeker, 2013; Sedivy, Tanenhaus, Chambers, & Carlson, 1999). Likewise, by age 2, children recruit subjecthood and repeated mention as cues to prominent entities and co-reference pronouns on this basis (Hartshorne et al., 2015; Pyykkönen et al., 2010, Arnold, Castro-Schilo, Zerkle, & Leela, 2019; Song & Fisher, 2005). By age 5, they exploit contrastive prosody to predict reference to new versus given entities (Arnold, 2008; Ito et al., 2012).

Finally, leveraging discourse to predict sentence meanings may offer crucial clues for learning words and structures during acquisition (Fisher et al., 2019; Horowitz & Frank, 2015; Sullivan & Barner, 2016). In pro-drop languages, salient arguments are often omitted from surface forms, creating challenges for isolating arguments associated with novel verbs. To solve this problem, Korean-learning 2-year-olds infer prominent entities from prior discourse (e.g., *What's Grandma doing?*) and deduce missing subjects in sentences (e.g., __ *was thomming the puppy* → infer NP1is *Grandma*, *thomming* is a transitive verb) (Fisher et al., 2019). Even in nonpro-drop languages like English, discourse continuity may be critical for interpreting the roughly one-third of child-directed utterances that omit subjects (Cameron-Faulkner, Lieven, & Tomasello, 2003; Laakso & Smith, 2007; Tardif, Shatz, & Naigles, 1997). Likewise, connections to developmental parsing may shed light on why imperatives are so prevalent when communicating with children (e.g., *Put on your socks!* → infer NP1 is *you*). While these constructions offer less information compared to declaratives, they move a reliable parsing cue (verbs) to sentence-initial position, and may promote more accurate predictions and less need for revision. This suggests one way in which caregiver input may be tailored to properties of developing parsing systems.

5.3. The development of syntactic parsing: Weather versus climate effects

The current findings offer insights into relationships between moment-to-moment comprehension and year-to-year development. In particular, they suggest that the inferences that generate the agent-first bias during development (i.e., learning form-to-meaning relations based on the distributions of NP1s, subjects, agents across many input sentences) are not the same as the ones governing its use during comprehension (i.e., identifying referents and parsing grammatical roles within a single sentence). This division between ontogenetic and chronometric processes explains how children can vary parsing strategies based on the uncertainty incurred by the communicative context. However, it highlights the difficulty of isolating pathways that mediate developmental and processing time scales. This problem is, in part, methodological. Even well-designed artificial language studies vastly simplify acquisition procedures in order to make them observable in 20-min tasks (see Zuhurudeen & Huang, 2016 for an alternative approach). Yet, this challenge also reflects theoretical limitations. It is difficult to imagine systematic pathways across two drastically different time scales.

To address the latter, we would like to offer an analogy: Processing is to development as weather is to climate. Weather tracks the momentary changes in the atmosphere, measured on a daily or hourly basis. It is influenced by temporary properties of the environment such as the jet stream, hurricanes, and clouds. Climate tracks the long-term averages of weather over years and decades. Unlike weather, it is shaped by stable properties in the environment like distance from the equator, proximity to oceans, and the earth's rotation around the sun. Climate affects weather in obvious ways. Due to their geography, Januarys in Honolulu, HI are far more pleasant than in College Park, MD. Moreover, human intervention can create unnatural but stable changes in the environment that alter climate (e.g., industrial revolution, dietary preferences), which in turn impacts weather (e.g., increased hurricanes, wildfires, flooding).

This analogy offers a framework for understanding relationships between processing and development. Like weather, processing is sensitive to momentary changes in communicative contexts, such as background noise and speaker-specific properties (e.g., listening in a bar, to an unfamiliar accent). The output of these procedures is sentence interpretation. Like climate, development reflects stable properties in learning environments, such as language(s), cultural background, and listener-specific properties (e.g., hearing English vs. Mandarin, family SES, using cochlear implants). These procedures yield lexicons and grammars. Development affects processing since 2-year-olds interpret sentences differently than 20-year-olds. Processing affects development since children learn linguistic regularities through iterative encounters with sentences. Yet, given massive differences in time scales, we hypothesize that processing only impacts development if it can induce stable changes in the learning environment. Our findings suggest two forms of "human intervention" of this kind. First, when children lack fine-grained knowledge to interpret all words in sentences, they may rely instead on broad-scale regularities within their language. This is the agent-first bias. Second, children may alter intake of input by privileging early-arriving cues over late-arriving ones. This is the kindergarten-path effect. While these dual strategies can generate comprehension errors, they may still yield optimal, long-term opportunities to bootstrap word meanings in probable structures (Huang & Arnold, 2016; Lidz et al., 2017; Reuter, Borovsky, & Lew-Williams, 2019).

This hypothesis contrasts with two dominant accounts in the literature. First, it has been argued that the kindergarten-path effect creates challenges for learning syntactic distinctions that are marked by late-arriving cues in sentences (Omaki & Lidz, 2015; Pozzan & Trueswell, 2015). For example, children's comprehension of causative verb morphology is less accurate in verb-final languages like Kannada (where morphemes are revision cues) compared to verb-initial languages like Tagalog (where they are prediction cues) (Lidz, Gleitman, & Gleitman, 2003; Trueswell, Kaufman, Hafri, & Lidz, 2012). English-speaking adults in an artificial-language task learn novel morphemes more successfully when they occur in a verb-initial language compared to a verb-final one (Pozzan & Trueswell, 2015). While this pathway shares similarities with our account above, our findings demonstrate that the kindergarten-path effect is only one element within a complex system for developmental parsing. In addition to within-sentence properties, children may leverage between-sentence discourse information to vary role assignment based on early- versus late-arriving parsing cues. This strategy solves basic challenges about how to determine who did what to whom, given varying sentence properties and linguistic knowledge. If the child is too conservative and eschews predictions altogether, they may fail to adequately interpret roles within fast-moving sentences. However, if the child is too liberal and generates predictions based solely on cue order rather than cue reliability, they may face widespread kindergarten-path challenges. By appealing to mechanisms that already exist for adults (Ferreira et al., 2002; Gibson et al., 2013; Levy et al., 2009), we argue that child parsing may be sensitive to uncertainty about linguistic properties (e.g., implications of new discourse referents) and may be more likely to adopt heuristics to interpret utterances in these contexts (e.g., agent-first bias).

Our hypothesis also differs from characterizations of processing and development as a single, undifferentiated procedure (Chang et al., 2006; Christiansen & Chater, 2016; Reuter, Emberson, Romberg, & Lew-Williams, 2018). Correlations between vocabulary size (a proxy for linguistic knowledge) and the speed of word recognition (Fernald, Marchman, & Weisleder, 2013; Marchman & Fernald, 2008), lexical prediction (Borovsky, Elman, & Fernald, 2012; Mani & Huettig, 2012), and syntactic revision (Anderson, Farmer, Goldstein, Schwade, & Spivey, 2011; Huang et al., 2017) are often used to support this perspective. Likewise, in syntactic-priming paradigms (Branigan & Messenger, 2016; Peter et al., 2015) and short-term training (Qi, Yuan, & Fisher, 2011; Yazbec, Kaschack, & Borovsky, 2019), children's predictions vary with the distributional statistics of recent input. This, in turn, can alter interpretation of syntactic structures, which influence bootstrapped meanings for novel words (Havron, de Carvalho, Fiévet, & Christophe, 2019). When exposed to sentence frames that prime noun interpretations, 3- and 4-year-olds infer object-related meanings but switch to event-related meanings when prior frames imply verbs instead. The authors conclude that, "*In the sense that processing changes children's expectations, it produces the same kind of effect that we would normally call learning, and there is no sense in talking about two different systems. Language acquisition is in fact a process by which the child is learning to process language (pg. 88).*"

However, this account does not specify the basis by which sentences are processed and linguistic knowledge is subsequently acquired (see Rabagliati, Gambi, & Pickering, 2016 for discussions of these issues). Moreover, given vast time-scale differences between processing and development, we should be careful not to conflate the two. Saying that "*acquisition is learning to process*" (Christiansen and Chater (2016), pg. 10) is true in some sense and not true in others, but it is unclear what it buys in terms of explanatory power without spelling out specific procedures that children engage in while interpreting sentences (e.g., syntactic parsing, coreferencing), its relations to specific sentence cues (e.g., word order, verb morphology), and how these relations scale across development. The current study suggests that even in the ostensibly "simple" case of passives, children's interpretation is mediated by complex interactions at multiple levels of interpretation. Perhaps relationships between processing and development are akin to how cultural evolution shapes the languages that children acquire while biological evolution constrains their capacity to do so. Given differences in time scales (thousands vs. millions of years), these procedures must be distinct and subject to different factors. Likewise, since individual sentences are a drop in the bucket of all experiences that children will leverage to learn language, each one may have mathematically small impacts on development.

6. Conclusion

The current study investigated children's strategies for role assignment, and focused on variable use of the agent-first bias when comprehending active and passive sentences. Our findings suggest that inferences of discourse status influence children's strategies for parsing initial arguments. Linguistic cues that signal new entities increase interpretive uncertainty, and this enhances the agent-first bias. In contrast, cues that imply given entities decrease uncertainty and enable children to delay role assignment until more information arrives later in sentences. These findings suggest that children simultaneously interpret relations between sentences (e.g., discourse continuity) and within (e.g., role assignment), such that commitments to the former can influence parsing cues for the latter. This work paves the way for future research examining the extent to which developmental parsing strategies enable inferences of word and sentence meanings and tracking of statistical regularities during language acquisition.

Note

1 We also conducted power analyses by examining correct responses by NP1 status, and estimating the sample sizes required for detecting construction differences with 80% power. When NP1s referred to new entities, we needed about 14 participants in Experiment 1 ($d = 0.71$), 6 participants in Experiment 2 ($d = 1.35$), and 5 participants in Experiment 3 (d=1.63) to distinguish actives and passives. In contrast, when NP1 referred to given entities, we would need 582 participants in Experiment 1 ($d = 0.10$), 73 participants in Experiment 2 ($d = 0.29$), and 38 participants in Experiment 3 ($d = 0.41$) to do so. This suggests the current study (with 20 participants in each NP1 status) was adequately powered to detect construction differences when children adopt an agent-first bias, but we would need far more participants to detect differences when the agent-first bias is not adopted.

Acknowledgments

This work has benefited from insightful conversations with Rochelle Newman, Jared Novick and Jeff Lidz. We would like to thank Lauren Abadie, Kerianna Frederick, Erin Hollister, Alison Arnold, and other research assistants in the Language and Cognition Laboratory for help in collecting and coding data. Portions of this work have been presented at the 2016 CUNY Conference on Human Sentence Processing. This work was supported by NSF-BCS grant BCS-1844194 (Huang).

Open Research Badges

This article has earned Open Data and Open Materials badges. Data and materials are available at https://osf.io/8shgf/ and materials are available at https://osf.io/8shgf/.

References

Abbot-Smith, K., Chang, F., Rowland, C., Ferguson, H., & Pine, J. (2017). Do two and three year old children use an incremental first-NP-as-agent bias to process active transitive and passive sentences?: A permutation analysis. *PloS One*, *12*(10). https://doi.org/10.1371/journal.pone.0186129

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. Arnon & 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.

Ariel, M. (1988). Referring and accessibility. *Journal of Linguistics*, *24*(1), 65–87.

Arnold, J. E. (1998). Reference form and discourse patterns. Ph.D., Stanford University.

Arnold, J. E. (2008). THE BACON not the bacon: How children and adults understand accented and unaccented noun phrases. *Cognition*, *108*, 69–99.

Arnold, J. E. (2010). How speakers refer: The role of accessibility. *Language and Linguistic Compass*, *4*, 187–203.

Arnold, J. E., Castro-Schilo, L., Zerkle, S., & Leela, R. A. O. (2019). Print exposure predicts pronoun comprehension strategies in children. *Journal of Child Language*, *46*(5), 863–893.

Arunachalam, S., & Waxman, S. R. (2011). Grammatical form and semantic context. *Language Learning and Development*, *7*, 169–184.

Arunachalam, S., & Waxman, S. R. (2015). Let's see a boy and a balloon: Argument labels and syntactic frame in verb learning. *Language Acquisition*, *22*, 117–131.

Arunachalam, S., Leddon, E. M., Song, H., Lee, Y., & Waxman, S. R. (2013). Doing more with less: Verb learning in Korean-acquiring 24-month-olds. *Language Acquisition*, *20*, 292–304.

Aston, G., & Burnard, L. (1998). *The BNC handbook: Exploring the British National Corpus with SARA*. Edinburgh, UK: Edinburgh University Press.

Baayen, R. H., Davidson, D. J., & Bates, D. M. (2008). Mixed-effects modeling with crossed random effects for subjects and items. *Journal of Memory and Language*, *59*, 390–412.

Barr, D. J., Levy, R., Scheepers, C., & Tily, H. J. (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.

Bates, E., & MacWhinney, B. (1981). Functionalist approaches to grammar. In E. Wanner & L. Gleitman (Eds.), *Language acquisition: The state of the art*. pp. 173–218 New York: Cambridge University Press.

Bates, E., MacWhinney, B., Caselli, C., Devescovi, A., Natale, F., & Venza, V. (1984). A cross-linguistic study of the development of sentence interpretation strategies. *Child Development*, *55*, 341–354.

Bates, E., McNew, S., MacWhinney, B., Devescovi, A., & Smith, S. (1982). Functional constraints on sentence processing: A cross-linguistic study. *Cognition*, *11*, 245–299.

Bever, T. G. (1970). The cognitive basis for linguistic structure. In J. R. Hayes (Ed.), *Cognition and the development of language* (pp. 279–362). New York: Wiley.

Bhattasali, S., & Resnik, P. (2020). Differentiating between broad & local context cues using surprisal: An fMRI study. Poster presented at the 33rd annual CUNY Human Sentence Processing Conference. Amherst, Massachusetts.

Bion, R. A., Borovsky, A., & Fernald, A. (2013). Fast mapping, slow learning: Disambiguation of novel word-object mappings in relation to vocabulary learning at 18, 24, and 30 months. *Cognition*, *126*, 39–53.

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.

Brandt, S., Kidd, E., Lieven, E., & Tomasello, M. (2009). The discourse bases of relativization: An investigation of young German and English-speaking children's comprehension of relative clauses. *Cognitive Linguistics*, *20*, 539–570.

Branigan, H. P., & Messenger, K. (2016). Consistent and cumulative effects of syntactic experience in children's sentence production: Evidence for error-based implicit learning. *Cognition*, *157*, 250–256.

Brown-Schmidt, S., & Tanenhaus, M. K. (2006). Watching the eyes when talking about size: An investigation of message formulation and utterance planning. *Journal of Memory and Language*, *54*, 592–609.

Brown-Schmidt, S., & Tanenhaus, M. (2008). Real-time investigation of referential domains in unscripted conversation: A targeted language game approach. *Cognitive Science*, *32*, 643–684.

Cameron-Faulkner, T., Lieven, E., & Tomasello, M. (2003). A construction based analysis of child directed speech. *Cognitive Science*, *27*, 843–873.

Chafe, W. (1987). Cognitive constraints on information flow. *Coherence and Grounding in Discourse*, *11*, 21–51.

Chan, A., Lieven, E., & Tomasello, M. (2009). Children's understanding of the agent-patient relations in the transitive construction: Cross-linguistic comparisons between Cantonese, German, and English. *Cognitive Linguistics*, *20*, 267–300.

Chang, F., Dell, G. S., & Bock, K. (2006). Becoming syntactic. *Psychological Review*, *113*, 234–272.

Choi, Y., & Trueswell, J. (2010). Children's (in)ability to recover from garden paths in a verb-final language: Evidence for developing control in sentence processing. *Journal of Experimental Child Psychology*, *106*, 41–61.

Christiansen, M. H., & Chater, N. (2016). The now-or-never bottleneck: A fundamental constraint on language. *Behavioral and Brain Sciences*, *39*, e62.

Dautriche, I., Cristia, A., Brusini, P., Yuan, S., Fisher, C., & Christophe, A. (2014). Toddlers default to canonical surface-to-meaning mapping when learning verbs. *Child Development*, *85*, 1168–1180.

De Villiers, J. G., & De Villiers, P. A. (1973). A cross-sectional study of the acquisition of grammatical morphemes in child speech. *Journal of Psycholinguistic Research*, *2*, 267–278.

Deen, K. U., Bondoc, I., Camp, A., Estioca, S., Hwang, H., Shin, G. H., Takahashi, M., Zenker, F., & Zhong, J. C. (2018). Repetition brings success: Revealing knowledge of the passive voice. *Proceedings of the 42nd annual boston university conference on language development* (pp. 200–213). In Anne B. Bertolini & Maxwell J. Kaplan (Eds.), Somerville, MA: Cascadilla Press.

Eager, C. D., & Roy, J. (2017*). Mixed models are sometimes terrible*. Austin, TX: Linguistic Society of America.

Fernald, A., Marchman, V. A., & Weisleder, A. (2013). SES differences in language processing skill and vocabulary are evident at 18 months. *Developmental Science*, *16*, 234–248.

Ferreira, F. (2003). The misinterpretation of noncanonical sentences. *Cognitive Psychology*, *47*, 164–203.

Ferreira, F., & Patson, N. D. (2007). The 'good enough' approach to language comprehension. *Language and Linguistics Compass*, *1*, 71–83.

Ferreira, F., Bailey, K. G., & Ferraro, V. (2002). Good-enough representations in language comprehension. *Current Directions in Psychological Science*, *11*, 11–15.

Ferreira, V. S., Slevc, L. R., & Rogers, E. S. (2005). How do speakers avoid ambiguous linguistic expressions? *Cognition*, *96*, 263–284.

Fisher, C., Jin, K. S., & Scott, R. M. (2019). The developmental origins of syntactic bootstrapping. *Topics in Cognitive Science*, *12*, 48–77.

Frank, S., Feldman, N., & Goldwater, S. (2014, June). Weak semantic context helps phonetic learning in a model of infant language acquisition. In *Proceedings of the 52nd Annual Meeting of the Association for Computational Linguistics* (Volume 1: Long Papers, pp. 1073–1083). Stroudsburg, PA: Association for Computational Linguistics.

Gibson, E., Bergen, L., & Piantadosi, S. T. (2013). Rational integration of noisy evidence and prior semantic expectations in sentence interpretation. *Proceedings of the National Academy of Sciences*, *110*, 8051–8056.

Givón, T. (1983). *Topic continuity in discourse: A quantitative cross-language study* (Vol. *3*). Amsterdam, The Netherlands: John Benjamins Publishing.

Gordon, P. C., Hendrick, R., & Johnson, M. (2004). Effects of noun phrase type on sentence complexity. *Journal of Memory and Language*, *51*, 97–114.

Gordon, P., & Chafetz, J. (1990). Verb-based versus class-based accounts of actionality effects in children's comprehension of passives. *Cognition*, *36*, 227–254.

Grice, H. P. (1975). Logic and conversation. In P. Cole & J. L. Morgan (Eds.), *Syntax and semantics* (Vol. *3*, pp. 42–58). New York: Academic Press.

Gries, S. (2015). The most under-used statistical method in corpus linguistics: Multi-level (and mixed-effects) models. *Corpora*, *10*, 95–125.

Grigoroglou, M., & Papafragou, A. (2019). Interactive contexts increase informativeness in children's referential communication. *Developmental Psychology*, *55*, 951–966.

Gundel, J. K., Hedberg, N., & Zacharski, R. (1993). Cognitive status and the form of referring expressions in discourse. *Language*, *69*, 274–307.

Hale, J. (2001). A probabilistic Earley parser as a psycholinguistic model. In *Proceedings of the second meeting of the north american chapter of the association for computational linguistics on language technologies* (pp. 1–8). Stroudsburg, PA: Association for Computational Linguistics.

Hartshorne, J. K., Nappa, R., & Snedeker, J. (2015). Development of the first-mention bias. *Journal of Child Language*, *42*, 423–446.

Havron, N., de Carvalho, A., Fiévet, A. C., & Christophe, A. (2019). Three-to four-year-old children rapidly adapt their predictions and use them to learn novel word meanings. *Child Development*, *90*, 82–90.

He, A. X., Kon, M., & Arunachalam, S. (2020). Linguistic context in verb learning: Less is sometimes more. *Language Learning and Development*, *16*, 22–42.

Hirsh-Pasek, K., & Golinkoff, R. M. (1996). *The origins of grammar: Evidence from early language comprehension*. Cambridge, MA: MIT Press.

Horowitz, A., & Frank, M. (2015). Young children's developing sensitivity to discourse continuity as a cue for inferring reference. *Journal of Experimental Child Psychology*, *129*, 84–97.

Horst, J. S., & Samuelson, L. K. (2008). Fast mapping but poor retention by 24-month-old infants. *Infancy*, *13*, 128–157.

Huang, Y., & Arnold, A. (2016). Word learning in linguistic context: Processing and memory effects. *Cognition*, *156*, 71–87.

Huang, Y., & Hollister, E. (2019). Developmental parsing and linguistic knowledge: Reexamining the role of cognitive control in the kindergarten path effect. *Journal of Experimental Child Psychology*, *184*, 210–219.

Huang, Y., & Snedeker, J. (2013). The use of lexical and referential cues in children's online interpretation of adjectives. *Developmental Psychology*, *49*, 1090–1102.

Huang, Y., & Snedeker, J. (2011). Processing and prediction in pragmatic inferencing: Understanding task-dependent effects in the generation of scalar implicatures. Poster presented at the 36th Boston University Conference on Language Development, November 4—6, Boston, MA.

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.

Imai, M., Haryu, E., & Okada, H. (2005). Mapping novel nouns and verbs onto dynamic action events: Are verb meanings easier to learn than noun meanings for Japanese children? *Child Development*, *76*, 340–355.

Imai, M., Li, L., Haryu, E., Okada, H., Hirsh-Pasek, K., Golinkoff, R. M., & Shigematsu, J. (2008). Novel noun and verb learning in Chinese-, English-, and Japanese-speaking children. *Child Development*, *79*, 979–1000.

Ito, K., Bibyk, S. A., Wagner, L., & Speer, S. R. (2014). Interpretation of contrastive pitch accent in six-to eleven-year-old English-speaking children (and adults). *Journal of Child Language*, *41*(1), 84–110.

Jaeger, T. F. (2008). Categorical data analysis: Away from ANOVAs (transformation or not) and towards logit mixed models. *Journal of Memory and Language*, *59*(4), 434–446.

Jaeger, T. F. (2009). Random effects: Should I stay or should I go? URL: Retrieved from https://hlplab.wordpress.com/2009/05/14/random-effect-structure/

Jarosz, A. F., & Wiley, J. (2014). What are the odds? A practical guide to computing and reporting Bayes factors. *The Journal of Problem Solving*, *7*, 2.

Johnson-Laird, P. N. (1968). The choice of the passive voice in a communicative task. *British Journal of Psychology*, *59*, 7–15.

Kidd, E., & Bavin, E. L. (2005). Lexical and referential cues to sentence interpretation: An investigation of children's interpretation of ambiguous sentences. *Journal of Child Language*, *32*(4), 855–876.

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.

Laakso, A., & Smith, L. B. (2007). Pronouns and verbs in adult speech to children: A corpus analysis. *Journal of Child Language*, *34*, 725–763.

Lee, M. D., & Wagenmakers, E. J. (2014). *Bayesian cognitive modeling: A practical course*. Cambridge, UK: Cambridge University Press.

Levy, R. (2008). Expectation-based syntactic comprehension. *Cognition*, *106*, 1126–1177.

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*, *106*, 21086–21090.

Lidz, J., Bunger, A., Leddon, E., Baier, R., & Waxman, S. R. (2010). When one cue is better than two: Lexical vs. syntactic cues to verb learning. Unpublished manuscript.

Lidz, J., Gleitman, H., & Gleitman, L. (2003). Understanding how input matters: Verb learning and the footprint of universal grammar. *Cognition*, *87*, 151–178.

Lidz, J., White, A. S., & Baier, R. (2017). The role of incremental parsing in syntactically conditioned word learning. *Cognitive Psychology*, *97*, 62–78.

MacWhinney, B., Bates, E., & Kliegl, R. (1984). Cue validity and sentence interpretation in English, German, and Italian. *Journal of Verbal Learning and Verbal Behavior*, *23*, 127–150.

MacWhinney, B., Pléh, C., & Bates, E. (1985). The development of sentence interpretation in Hungarian. *Cognitive Psychology*, *17*, 178–209.

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.

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

Markman, E. M., & Wachtel, G. F. (1988). Children's use of mutual exclusivity to constrain the meanings of words. *Cognitive Psychology*, *20*, 121–157.

Martin, I., Goupell, M. & Huang, Y. (in press). Children's syntactic parsing and sentence comprehension with a degraded auditory signal. To appear in The Journal of the Acoustical Society of America.

Messenger, K., & Fisher, C. (2018). Mistakes weren't made: Three-year-olds' comprehension of novel-verb passives provides evidence for early abstract syntax. *Cognition*, *178*, 118–132.

Moineddin, R., Matheson, F. I., & Glazier, R. H. (2007). A simulation study of sample size for multilevel logistic regression models. *BMC Medical Research Methodology*, *7*(1), 1.

Montgomery, J. W., Gillam, R. B., Evans, J. L., & Sergeev, A. V. (2017). "Whatdunit?" Sentence comprehension abilities of children with SLI: Sensitivity to word order in canonical and noncanonical structures. *Journal of Speech, Language, and Hearing Research*, *60*, 2603–2618.

Morey, R. D., & Rouder, J. N. (2011). Bayes factor approaches for testing interval null hypotheses. *Psychological Methods*, *16*, 406.

Omaki, A., & Lidz, J. (2015). Linking parser development to acquisition of syntactic knowledge. *Language Acquisition*, *22*, 158–192.

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.

Oppenheimer, K., Morini, G., Strother-Garcia, K., Huang, Y., & Owen Van Horne, A. (2020). Online comprehension of passives by children with Developmental Language Disorder and typically developing age-matched peers. Poster presented at the 45th Boston University Conference on Language Development. Boston, MA.

Ovans, Z., Huang, Y., & Feldman, N. (2020). The (un)surprising kindergarten path. In S. Denison, M. Mack, Y. Xu, & B. C. Armstrong (Eds.), *Proceedings of the 42nd annual conference of the cognitive science society* (pp. 350–356). Austin, TX: Cognitive Science Society.

Ovans, Z., Novick, J., & Huang, Y. (2020). Engaging cognitive control may help children ignore unreliable cues during sentence processing. Poster presented at the 33st Annual CUNY Conference on Human Sentence Processing. Amherst, MA.

Perkins, L., Feldman, N., & Lidz, J. (2017). April). Learning an input filter for argument structure acquisition. In Proceedings of the 7th Workshop on Cognitive Modeling and Computational Linguistics (CMCL 2017) (pp. 11–19).

Peter, M., Chang, F., Pine, J. M., Blything, R., & Rowland, C. F. (2015). When and how do children develop knowledge of verb argument structure? Evidence from verb bias effects in a structural priming task. *Journal of Memory and Language*, *81*, 1–15.

Pozzan, L., & Trueswell, J. C. (2015). Revise and resubmit: How real-time parsing limitations influence grammar acquisition. *Cognitive Psychology*, *80*, 73–108.

Pyykkönen, P., Matthews, D., & Järvikivi, J. (2010). Three-year-olds are sensitive to semantic prominence during online language comprehension: A visual world study of pronoun resolution. *Language and Cognitive Processes*, *25*, 115–129.

Qi, Z., Yuan, S., & Fisher, C. (2011). Where does verb bias come from? Experience with particular verbs affect online sentence processing. *In Proceedings of the 35th Boston university conference on language development* (pp. 500–512). Somerville, MA: Cascadilla Press.

Rabagliati, H., Gambi, C., & Pickering, M. J. (2016). Learning to predict or predicting to learn? *Language, Cognition and Neuroscience*, *31*, 94–105.

Reuter, T., Borovsky, A., & Lew-Williams, C. (2019). Predict and redirect: Prediction errors support children's word learning. *Developmental Psychology*, *55*, 1656–1166.

Reuter, T., Emberson, L., Romberg, A., & Lew-Williams, C. (2018). Individual differences in nonverbal prediction and vocabulary size in infancy. *Cognition*, *176*, 215–219.

Rohde, H., & Frank, M. C. (2014). Markers of topical discourse in child-directed speech. *Cognitive Science*, *38*, 1634–1661.

Roy, B. C., Frank, M. C., DeCamp, P., Miller, M., & Roy, D. (2015). Predicting the birth of a spoken word. *Proceedings of the National Academy of Sciences*, *112*(41), 12663–12668.

Sedivy, J., Tanenhaus, M., Chambers, C., & Carlson, G. (1999). Achieving incremental semantic interpretation through contextual representation. *Cognition*, *71*, 109–147.

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

Song, H., & Fisher, C. (2005). Who's "she"? Discourse prominence influences preschoolers' comprehension of pronouns. *Journal of Memory and Language*, *52*, 29–57.

Sullivan, J., & Barner, D. (2016). Discourse bootstrapping: Preschoolers use linguistic discourse to learn new words. *Developmental Science*, *19*, 63–75.

Tardif, T., Shatz, M., & Naigles, L. (1997). Caregiver speech and children's use of nouns versus verbs: A comparison of English, Italian, and Mandarin. *Journal of Child Language*, *24*, 535–565.

Trueswell, J. C., Kaufman, D., Hafri, A., & Lidz, J. (2012). Development of parsing abilities interacts with grammar learning: Evidence from Tagalog and Kannada. In *Proceedings of the 36th annual Boston university conference on language development* (pp. 620–632). Somerville, MA: Cascadilla Press.

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.

van Doorn, J., Ly, A., Marsman, M., & Wagenmakers, E. J. (2020). Bayesian rank-based hypothesis testing for the rank sum test, the signed rank test, and Spearman's ρ. *Journal of Applied Statistics*, *47*, 2984–3006.

Warren, T., & Gibson, E. (2002). The influence of referential processing on sentence complexity. *Cognition*, *85*(1), 79–112.

Waxman, S. R., Lidz, J. L., Braun, I. E., & Lavin, T. (2009). Twenty four-month-old infants' interpretations of novel verbs and nouns in dynamic scenes. *Cognitive Psychology*, *59*(1), 67–95.

Weighall, A. (2008). The kindergarten path effect revisited: Children's use of context in processing structural ambiguities. *Journal of Experimental Child Psychology*, *99*, 75–95.

Williams, E. S. (1977). Discourse and logical form. *Linguistic Inquiry*, 101–139.

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.

Yazbec, A., Kaschak, M. P., & Borovsky, A. (2019). Developmental timescale of rapid adaptation to conflicting cues in real-time sentence processing. *Cognitive Science*, *43*, e12704.

Zuhurudeen, F., & Huang, Y. (2016). Effects of statistical learning on the acquisition of grammatical categories through Qur'anic memorization: A natural experiment. *Cognition*, *148*, 79–84.

Appendix A

List of critical items for Experiment 1

Condition	Critical sentence (active/passive)	Mentioned object	Likely agent	Likely theme
Definite NP1	The boy is gently kicking/kicked by it	BOY	HORSE	BALL
Pronoun NP1	It is gently kicking/kicked by the boy			
Definite NP1	The towel is gently cleaning/cleaned by it	TOWEL	DRYER	PAN
Pronoun NP1	It is gently cleaning/cleaned by the towel			
Definite NP1	The firefighter is quickly rescuing/rescued by it	FIREFIGHTER	HELICOPTER	CAT
Pronoun NP1	It is quickly rescuing/rescued by the firefighter			
Definite NP1	The girl is happily feeding/fed by her	GIRL	MOTHER	BABY
Pronoun NP1	She is happily feeding/fed by the girl			
Definite NP1	The rabbit is slowly eating/eaten by it	RABBIT	WOLF	CARROTS
Pronoun NP1	It is slowly eating/eaten by the rabbit			
Definite NP1	The seal is quickly eating/eaten by it	SEAL	SHARK	FISH
Pronoun NP1	It is quickly eating/eaten by the seal			
Definite NP1	The dog is slowly chasing/chased by it	DOG	CAR	RABBIT
Pronoun NP1	It is slowly chasing/chased by the dog			
Definite NP1	The girl is tightly hugging/hugged by her	GIRL	MOTHER	BABY
Pronoun NP1	She is tightly hugging/hugged by the girl			
Definite NP1	The frog is quietly catching/caught by it	FROG	DOG	FLY
Pronoun NP1	It is quietly catching/caught by the frog			
Definite NP1	The boy is carefully lifting/lifted up by him	BOY	DAD	BABY
Pronoun NP1	He is carefully lifting/lifted up by the child			
Definite NP1	The rock is loudly smashing/smashed by it	ROCK	HAMMER	EGG
Pronoun NP1	It is loudly smashing/smashed by the rock			
Definite NP1	The cat is easily scaring/scared by it	CAT	DOG	MOUSE
Pronoun NP1	It is easily scaring/scared by the cat			

Appendix B

List of critical items for Experiment 2

Condition	Prime	Critical sentence (active/passive)	Mentioned object	Competitor	Likely agent	Likely theme
Conjoined	The brother and the boy dance	He is carefully lifting up/lifted up by the boy	BOY	BROTHER	DAD	DOLL
Single	The brother dances					
Conjoined	The purple brush and the pink brush fall	It is gently washing/washed by the pink brush	PINK BRUSH	PURPLE BRUSH	TOWEL	PAN
Single	The purple brush falls					
Conjoined	The firefighter and the policeman laugh	He is quickly rescuing/rescued by the policeman	POLICEMAN	FIREFIGHTER	SOLDIER	BOY
Single	The firefighter laughs					
Conjoined	The nurse and the patient walk	She is slowly holding/held by the patient	PATIENT	NURSE	DOCTOR	BABY
Single	The nurse walks					
Conjoined	The purple rabbit and the white rabbit run	It is gently kicking/kicked by the white rabbit	WHITE RABBIT	PURPLE RABBIT	WOLF	BUTTERFLY
Single	The purple rabbit runs					
Conjoined	The gray seal and the white seal swim	It is suddenly catching/caught by the white seal	WHITE SEAL	GRAY SEAL	SHARK	FISH
Single	The gray seal swims					
Conjoined	The friend and the boy laugh	He is quickly kicking/kicked by the boy	BOY	FRIEND	BULLY	BROTHER
Single	The friend laughs					
Conjoined	The mom and the aunt sing	She is tightly hugging/hugged by the aunt	AUNT	MOM	GRANDMA	GIRL
Single	The mom sings					
Conjoined	The green frog and the orange frog sleep	It is gently pushing/pushed by the orange frog	ORANGE FROG	GREEN FROG	DOG	FLY
Single	The green frog sleeps					
Conjoined	The sister and the girl dance	She is gently spinning/spun by the girl	GIRL	SISTER	MOTHER	DOLL
Single	The sister dances					
Conjoined	The orange ball and the blue ball roll	It is loudly smashing/smashed by the blue ball	BLUE BALL	ORANGE BALL	HAMMER	EGG
Single	The orange ball rolls					
Conjoined	The white cat and the black cat eat	It is quickly scaring/scared by the black cat	BLACK CAT	WHITE CAT	DOG	MOUSE
Single	The white cat eats					

Appendix C

List of critical items for Experiment 3

Condition	Critical Sentence (active/passive)	Familiar Item	Novel Agent	Novel Theme
Known NP1	The car will be loudly squishing/squished by the noytoff			
Novel NP1	The vaychip will be quickly grabbing/grabbed by the mouse	MOUSE	VULTURE-LIKE CREATURE	PLANT-LIKE OBJECT
Known NP1	The mouse will be quickly grabbing/grabbed by the vaychip			
Novel NP1	The bellwer will be quickly chasing/chased by the fox	FOX	OCTOPUS-LIKE CREATURE	RODENT-LIKE CREATURE
Known NP1	The fox will be quickly chasing/chased by the bellwer			
Novel NP1	The leepo will be slowly eating/eaten by the rabbit	RABBIT	WOLF-LIKE CREATURE	CABBAGE-LIKE PLANT
Known NP1	The rabbit will be slowly eating/eaten by the leepo			
Novel NP1	The blicket will be quickly chasing/chased by the seal	SEAL	SEA MONSTER	PUNY CREATURE
Known NP1	The seal will be quickly chasing/chased by the blicket			
Novel NP1	The coopa will be quickly chasing/chased by the dog	DOG	CARRIAGE-LIKE VEHICLE	RABBIT-LIKE CREATURE
Known NP1	The dog will be quickly chasing/chased by the coopa			
Novel NP1	The furpin will be quietly scared by the monkey	MONKEY	ROBOT-LIKE CREATURE	BUG-LIKE CREATURE
Known NP1	The monkey will be tightly hugging/hugged by the furpin			
Novel NP1	The daylon will be quietly catching/caught by the frog	FROG	DOG-LIKE CREATURE	FLY-LIKE CREATURE
Known NP1	The frog will be quietly catching/caught by the daylon			
Novel NP1	The chowvag will be carefully lifting/lifted up by the girl	GIRL	CENATAUR	SPOTTY BOX
Known NP1	The girl will be carefully lifting/lifted up by the chowvag			
Novel NP1	The tayvak will be loudly smashing/smashed by the rock	ROCK	HAMMER-LIKE OBJECT	CONE-LIKE OBJECT
Known NP1	The rock will be loudly smashing/smashed by the tayvak			
Novel NP1	The nedoke will be easily scaring/scared by the cat	CAT	DOG-LIKE CREATURE	MOUSE-LIKE CREATURE
Known NP1	The cat will be easily scaring/scared by the nedoke			