Journal of Child Language (2024), **51**, 271–287
doi:10.1017/S0305000922000678

CAMBRIDGE UNIVERSITY PRESS

ARTICLE

Inhibitory Control and Patterns of Errors in Resolution of Syntactically Ambiguous Sentences

Kaitlyn E. MAY[1,*] and Jason SCOFIELD[2]

[1]Department of Educational Studies in Psychology, Research Methodology, and Counseling, The University of Alabama, USA.
[2]Department of Human Development and Family Studies, The University of Alabama, USA.
*Corresponding author. Kaitlyn E. May, Department of Educational Studies in Psychology, Research Methodology, and Counseling, University of Alabama, Tuscaloosa, AL 35487-0160. Email: kmay3@crimson.ua.edu.

(Received 26 October 2021; revised 06 June 2022; accepted 06 November 2022)

Abstract

Sentences that have more than one possible meaning are said to be syntactically ambiguous (SA). Because the correct interpretation of these sentences can be unclear, resolving SA sentences can be cognitively demanding for children, particularly with regards to inhibitory control (IC). In this study we provide three lines of evidence supporting the importance of IC in SA resolution. First, we show that children with higher IC resolve more SA sentences correctly. Second, we show that SA resolution is worse on tasks that place higher demands on IC, even for children with high IC. Third, we show that children with higher IC make different types of SA errors than children with lower IC. This study expands understanding of the cognitive skills underlying language and suggests a need to consider task demands on IC when developing educational curriculums.

Keywords: syntactic ambiguity; inhibitory control; executive functions; sentence processing; child development

Introduction

Listeners process sentences in real-time. This means that any parsing and revision needed to interpret the sentence must happen in the moment, as the sentence is being heard (Woodard, Pozzan, & Trueswell, 2016). Because of the nature of real-time processing, it is possible for listeners to encounter words or phrases that can temporarily ambiguate the meaning of a sentence. Thus, while in-the-moment processing can be efficient, it can potentially come at a cost of accuracy in cases where a word has more than one meaning or a phrase could be interpreted in multiple ways. For instance, consider the example depicted in Figure 1. Here, the sentence states “Put the frog on the napkin on the box.” As the listener processes the sentence in real-time, they may first interpret the sentence as “Put the frog on the napkin” and therefore have to update their interpretation as additional information is provided (e.g., “on the box”). For listeners who must revise the initial interpretation, the meaning of the full sentence is ambiguous because the phrase

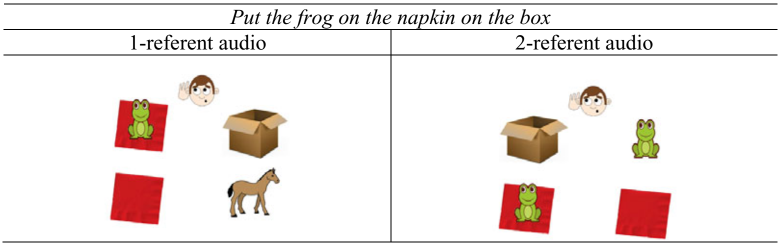

Figure 1. SA Task Sample Questions

"on the napkin" could be interpreted as referring to either the initial location of the frog (i.e., the frog on the napkin) or instructions for where to move the frog (i.e., to the napkin on the box). These are referred to as syntactically ambiguous (SA) sentences because they are structured in such a way that the listener's initial interpretation (in this case, to put the frog on the napkin) is more than likely to be incorrect and must later be resolved. This ambiguity only increases in tasks that feature two frogs, one of which is sitting on a napkin.

SA resolution has been studied in both children and adults. When faced with the example sentence in Figure 1, adults are able to revise their initial interpretation that the frog should be placed on the napkin, correcting their interpretation to move the frog *that's* on the napkin onto the box (Choi & Trueswell, 2010; Eberhard, Spivey-Knowlton, Sedivy, & Tanenhaus, 1995; Macdonald, Pearlmutter, & Seidenberg, 1994; Tanenhaus, Spivey-Knowlton, Eberhard, & Sedivy, 1995; Weighall, 2008). Children, however, are particularly burdened by SA – so much so that children's difficulty with SA is referred to as the 'kindergarten-path effect' (Anderson, Farmer, Goldstein, Schwade, & Spivey, 2011; Choi & Trueswell, 2010; Hurewitz, Brown-Schmidt, Thorpe, Gleitman, & Trueswell, 2000; Kidd & Bavin, 2005; Trueswell & Gleitman, 2004; Trueswell, Sekerina, Hill, & Logrip, 1999; Weighall, 2008). When presented with the aforementioned sentence, young children typically commit to their first interpretation, placing the frog on the empty napkin even after hearing the remainder of the sentence (Anderson et al., 2011; Choi & Trueswell, 2010; Christianson, Hollingworth, Halliwell, & Ferreira, 2001; Hurewitz et al., 2000; Spivey, Tanenhaus, Eberhard, & Sedivy, 2002; Weighall, 2008; Woodard et al., 2016). Regardless of the number of referents, children tend to prefer to interpret "on the napkin" as directive of where to move the frog rather than the description of which frog to move. Moreover, children's choice of which frog to move in the 2-referent condition is often at chance and correlated to which object the child looks towards first (Trueswell et al., 1999). Conversely, children respond to unambiguous sentences of similar complexity (e.g., "Put the frog *that's* on the napkin on the box"), with near-perfect accuracy (Woodard et al., 2016). This suggests that children have difficulty with SA sentences not because they are complex, but because they are ambiguous.

Follow up studies replicated the kindergarten-path effect while also showing improved processing of, and sensitivity to, SA with development (Anderson et al., 2011; Choi & Trueswell, 2010; Christianson et al., 2001; Hurewitz et al., 2000; Spivey et al., 2002; Weighall, 2008; Woodard et al., 2016). However, children begin to comprehend and use different sources of ambiguous information at different ages. By age 3, children can use gestures to interpret lexical ambiguity (Kidd & Holler, 2009). As early as 4- or 5-years of age, children are able to interpret verb biased SA sentences (e.g., "tickle the pig with the

https://doi.org/10.1017/S0305000922000678 Published online by Cambridge University Press

fan" or "surprise the pig with the fan," where tickle is more likely to reference an instrument use than surprise) using the probability that a verb will appear with one meaning versus another (Snedeker & Trueswell, 2004). However, 4-5-year-olds are not able to integrate context to interpret SA sentence: for instance, whether there were two pigs – one with a fan, and one without – in the picture. Moreover, even 8-12 year-olds struggle to use semantic plausibility to understand SA sentences (Traxler, Morris, & Seely, 2002).

A consensus within this body of work is that children's difficulty with SA reflects a difficulty to override their initial interpretation. If this is true, then perhaps children's failure to revise their initial interpretation is indicative of still-developing domain-general cognitive skills that support the parsing system, such as inhibitory control (IC). The present study examines the role of IC in resolving SA by assessing the presence and frequency of specific error types in children's SA resolution. We focused on error rates and error types because although it is well documented that children are prone to errors during SA tasks, error types have not been a major focus of the SA literature. Further examination of SA error patterns could be especially informative when considering a role for IC.

Inhibitory control and syntactically ambiguous sentences

Language processing requires both cognitive and linguistic abilities. Executive functions (EF), a set of cognitive processes critical to cognitive control of behavior and goal-directed thought and action, is thought to be one of these capacities supporting language processing. EF are often divided into three factors: inhibitory control (IC), working memory (WM), and cognitive flexibility (CF) (Diamond, 2013) and are critical to exerting control over thoughts and acts in accordance with task demands (Knight & Stuss, 2002). We were particularly interested in whether children's emerging IC contributes to SA performance, as IC is one of the earliest components of EF to develop and may be especially relevant to successfully resolving SA sentences by supporting inhibition of the initial interpretation.

Children's errors when resolving SA sentences indicate a difficulty to revise their interpretation in line with incoming information. Trueswell et al. (1999) initially argued that this inability to revise stems from limited processing resources. Subsequent work suggested a potential role for EF (Novick, Trueswell, & Thompson-Schill, 2005; Trueswell, Papafragou, & Choi, 2011). For example, it has been suggested that children's difficulty with SA is due to preemptive planning of a response and an inability to inhibit this plan in the face of new information (Meroni & Crain, 2003). In line with this work, it may be that children plan a response before they have received all information, consequently interpreting the words and phrases in the sentence in the order they are heard. The inability to inhibit the initial plan may reflect underdeveloped inhibitory control (IC) in which children struggle significantly to override automatic interpretations, failing to inhibit their prepotent response. Therefore, it is plausible that children's developing IC play a complex role in their ability to resolve SA, with more need for specific components of EF than others.

While few studies have directly assessed the role of EF in SA performance, particularly in childhood, there is some evidence from adults that suggests it might be important (e.g., Novick et al., 2005). Behavioral evidence likewise suggests a role for EF in SA processing, such that EF might mediate SA resolution (Choi & Trueswell, 2010; Trueswell et al., 1999;

Weighall, 2008; Woodard et al., 2016) and that EF is important to interpreting ambiguous stimuli more broadly (Wimmer, Doherty, & Collins, 2011). EF training has been shown to improve SA resolution in adults (Hussey & Novick, 2012; Novick, Hussey, Teubner-Rhodes, Harbison, & Bunting, 2013) and priming individuals to use EF in an incongruent Stroop task accelerates SA resolution response time (Hsu & Novick, 2016). A recent study demonstrated that EF predicted how many action errors (for example, first placing the frog on the empty napkin, then hopping it to the box) children made in processing SA sentences (Qi, Love, Fisher, & Brown-Schmidt, 2020). The authors argued that children with better EF may be better equipped to inhibit their initial interpretation in order to detect the ambiguity, pointing to the importance of IC specifically. Moreover, higher IC may support the ability to correct the first interpretation. However, Qi et al. (2020) exclusively studied EF and SA in 5-year-olds, warranting additional research in a wider age range across the preschool years. Taken together these studies provide support for the notion that SA resolution implicates IC.

Errors in SA performance

Most SA studies focus on accuracy when resolving sentences and, as previously mentioned, show that children tend to have poor accuracy (54% in Choi & Trueswell, 2010; 60% in Trueswell et al., 1999). Because children typically commit to their early interpretation, the specific error that causes their inaccuracy is not always clear. Few studies have explored the types of errors that children make on SA tasks. However, the types of errors and the frequency with which children make them during SA tasks may be helpful for identifying the factors underlying children's difficulty with SA. In their seminal paper, Trueswell et al. (1999) were able to document and discuss the types of errors frequently committed by children when resolving SA, although error type was not used as an outcome measure in their analyses.

Among the errors Trueswell et al. observed were object errors and destination errors. Object errors might occur if children's misinterpretation occurs as a result of attending mostly to the information presented later in the sentence (e.g., "…in the box"). In contrast, destination errors might occur if children's misinterpretation occurs as a result of attending mostly to the information presented early in the sentence (e.g., "put the frog on the napkin…"). Object and destination errors are interesting and potentially informative in the context of IC because they implicate its core subcomponents (Kim & Christianson, 2013; Woodard et al., 2016). Because IC is generally characterized by an ability to inhibit a prepotent response, and poor performance on SA tasks is commonly due to a failure to override the first interpretation (Choi & Trueswell, 2010; Engelhardt, Nigg, & Ferreira, 2017; Trueswell et al., 1999), it may be that IC plays a key role in SA via an individual's propensity to commit certain types of errors.

Current study

In the current study, we examine the role of IC in resolving SA. In doing so, we extend Trueswell et al.'s observations by directly examining the presence and frequency of object and destination errors in children's resolution of SA sentences. We focused on error rates and error types because young children are prone to commit errors during SA tasks and we believe that the types of errors committed by children with varying IC might differ, largely due to a tendency to perseverate and increased difficulty revising the first interpretation.

We hypothesize that children with lower IC would struggle to revise the first interpretation. In the 1-referent condition, if the child were to stop at the initial interpretation, they would put the frog on the napkin. This would result in a destination error because they did not listen to the remaining prompt and acted preemptively, placing the frog on the napkin rather than on the box, which would be the correct interpretation. Conversely, in the 2-referent condition this might result in the child moving the lone frog rather than the frog sitting on the napkin (an object error). Therefore, we hypothesize that children with lower IC will be more prone to destination errors in the 1-referent condition and to object errors in the 2-referent condition. Thus, the major goal of this study is to investigate the relationship between IC, error types, and SA performance in children aged 5- to 7 years.

Method

Participants

Seventy-five children ranging from 5 to 7 years old participated in this study ($M = 6.00$). There were 25 5-year-olds, 25 6-year-olds, and 25 7-year-olds. This age range was selected because syntactic ambiguity resolution skills emerge around five years in the literature. We recruited participants with parental consent and participant assent from local elementary schools. The participants represented a variety of socioeconomic statuses (SES), ranging from lower-income schools where the majority of students are on free or reduced lunch ($M = 77\%$) to a private school with predominantly upper-middle class/ upper class families.

Study design

This study had three main independent variables: referent, IC, and error type. Referent was a categorical, within-subjects variable. Referent was manipulated across two levels: 1-referent or 2-referent. The present study includes both 1-referent and 2-referent conditions, as the likelihood of making an error changes depending on the type of task. Specifically, in the 1-referent condition, destination errors are far more likely to occur whereas in the 2-referent condition it is possible to commit an object or destination error. Error type (object vs. destination) was a categorical within-subjects variable. Object errors were any trial in which the participant selected the incorrect object to move, whereas destination errors were any trial in which the object was placed in the incorrect destination. IC was a continuous between-subjects variable derived from the Rasch analyses described in the Results section. For the SA task, participants completed 14 trial in each referent condition (1-referent vs. 2-referent) for a total of 28 trials. Each trial featured a new SA sentence and a new display of objects (see Supplementary Materials).

Procedure

Consented and assented children were tested in a quiet, familiar place at their school. They were told that they would play a series of computer games with the experimenter and that these games would involve picture making (Syntactic Ambiguity Task), being silly (Day-Night Stroop), and feeding fish (Flanker Task). They were also told that there were no right or wrong answers. All tasks were computerized and presented on a touch-screen laptop using SuperLab Version 5.0.5 for Windows. Participants completed the tasks using the touch-screen and were monitored by a trained experimenter. If a child did not pass the training trials of any task, data collection immediately concluded. After completing the tasks, children were thanked for their participation, given a sticker, and escorted back to class. All tasks were presented in a fixed order, outlined below, without breaks.

Syntactic ambiguity task

Children completed a computerized syntactic ambiguity task modeled after the physical visual world task described in Trueswell et al. (1999), consisting of 1-referent and 2-referent trials. Because the task is a computerized replication of the task described in Trueswell et al. (1999), filler sentences were not used. Instead, each trial presented the child with an SA sentence. To mitigate the effects of multiple trials and of trial order, 7 counterbalanced versions of the task were used.

The task began with 4 practice trials of two objects with simple commands (i.e., "Put the ball in the basket.") to ensure the child understood the task. After four practice trials children proceeded to complete 28 randomized test trials consisting of 14 1-referent and 14 2-referent trials. Each trial type involved two parts. First, children saw a set of four images displayed on the touch-screen laptop. The set included two object images (e.g., a frog on a napkin and a horse for the 1-referent, a frog on a napkin and another frog for the 2-referent) and two destination images (e.g., an empty box and an empty napkin). Second, children heard a SA sentence via an audio recording (e.g., "Put the frog on the napkin on the box"). In 1-referent trials, there was one image of the subject of the sentence; while in the 2-referent sentence there were two images of the subject, thus increasing ambiguity. Children were asked to accurately illustrate the sentence and "put" the correct object on the correct destination by using the touch screen to drag and drop the object to its proper location. Each trial had an image of a boy with hand to ear at the top of the screen to remind the child to listen. See Figure 1. This task took approximately 15 minutes to complete.

Day-Night Stroop

To investigate children's ability to override an automatic response, children completed a modified, computerized version of the Day-Night Stroop (Gerstadt, Hong, & Diamond, 1994). The present study included two Stroop pairs: day versus night and girl versus boy (Gerstadt et al., 1994; Livesey, Keen, Rouse, & White, 2006) to increase the difficulty of the task due to initial data indicating that participants in the sample had near-perfect accuracy on the traditional Day-Night Stroop task. This data was excluded from analyses and the Boy-Girl stimuli were then added to increase the ability of the task to differentiate between IC in our sample.

The participant was presented with a picture of a moon and instructed, "When you see this picture, I want you to say 'day." The experimenter then asked the participant, "Can you say day?" This process was repeated with a picture of a sun ("say 'night'"), a picture of a boy ("say 'girl'"), and a picture of a girl ("say 'boy'"). The participant then completed a set of 4 practice trials randomized with 1 presentation of each picture. If the child responded correctly to practice trials, the experimenter praised the child and proceeded to testing. If the child responded incorrectly to the practice trials, the experimenter immediately started over, beginning with the instructions. If children failed to name each object

Figure 2. Day-Night Stroop Task

by its antonym during training, the experimenter would provide the antonym and ask the child to repeat it. Children did not receive feedback on test trials. Presentation order was randomized for each child. If the child demonstrated an understanding of the task, they proceeded to test trials and completed 32 trials randomized with 8 of each card. During test trials the child received no prompts. If the child hesitated, the experimenter prompted the child by saying, "What do you say for this one?" but never said the words day, night, boy, or girl. Each of the 32 trials were scored dichotomously (0 = *Incorrect*, 1 = *Correct*). This task took approximately 5 minutes to complete. See Figure 2.

Flanker task

To examine children's ability to disregard irrelevant information, inhibit, and to switch between different trial types, children completed a computerized Flanker task designed to replicate the NIH Toolbox Attentional Network Task. Children were told that they were going to play a game where they would "feed the fish." Fish appeared in a row in the center of the screen and children were instructed to feed the middle fish by clicking the arrow that matched the direction that the fish was pointing. On congruent trials the flanking fish were pointing in the same direction as the middle fish. In half of the congruent trials the fish were pointing to the right; in half they were pointing to the left. On incongruent trials the flanking fish were pointing in the opposite direction of the middle fish. In half of these the middle fish was pointing to the right with flanking fish pointing to the left; this was reversed for the other half of incongruent trials.

Children completed 4 randomized training trials consisting of 2 congruent trials (1 right and 1 left) and 2 incongruent trials (1 right and 1 left). If children were successful on the training trials they completed 20 randomized trials of 10 congruent (5 left, 5 right) and 10 incongruent trials (5 each direction) lasting 1700 ms or until the child responded. Fixations were randomized to occur between 400 and 1600 ms. Each trial was scored dichotomously (0 = *Incorrect*, 1 = *Correct*) and the 20 trials were then averaged to produce an overall accuracy score. This task took approximately 5 minutes to complete. See Figure 3.

https://doi.org/10.1017/S0305000922000678 Published online by Cambridge University Press

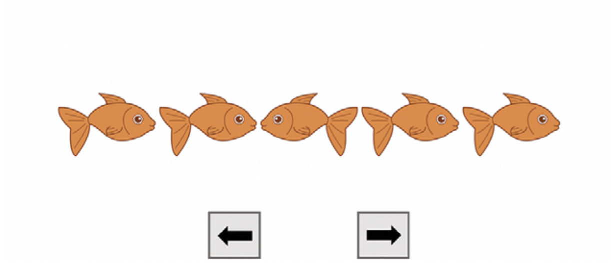

Figure 3. Flanker Task

Results

Rasch analysis of inhibitory control tasks

The dichotomous Rasch model (Rasch, 1960) was used to compute a latent IC variable for each participant, based on performance on the 2 IC tasks. First, an analysis of each of the IC tasks was conducted using the dichotomous Rasch model. The results of each model were reviewed and compiled into a single dataset and analyzed using the dichotomous Rasch model to gather evidence related to the validity of the interpretation and use of scores as a single, latent EF variable. Notably, the consolidated model performed better psychometrically than the individual models. Overall, this allowed for items from all tasks and respondents to be estimated on the same dimension and become directly comparable. This technique allows for testing of unidimensionality, takes into account response options which may not be equally spaced, and is generalizable across samples and items. The results of this research provide proof-in-concept of an innovative approach to measure EF that allows for consistency in assessment, addresses issues of task sensitivity, and facilitates conclusions across individual tasks and studies.

The dichotomous Rasch model uses sum scores from ordinal responses to calculate interval-level estimates that represent person locations (i.e., person ability) and item locations (i.e., the difficulty to provide a correct or positive response) on a linear scale that represents the latent variable (the log-odds or logit scale). The analysis was conducted using the "TAM" package for R (Robitzsch & Lüdtke, 2020) using the dichotomous Rasch Model equation:

$$ln\left[\frac{\varnothing_{ni1}}{\varnothing_{ni0}}\right] = \theta_n - \delta_i$$

The dichotomous Rasch model was selected among potential latent trait models because the DNS and Flanker tasks are scored dichotomously (0 = *Incorrect*, 1 = *Correct*). For each task item, the proportion of correct responses, logit-scale calibration (δ), *SE*, and model-data fit statistics were examined to determine the most difficult and easiest items. The proportion of correct responses, their logit-scale measure (θ), *SE*, and model-data fit statistics were examined to assess participant ability. Finally, the calibrations of the participants and items on the logit scale that represents the latent variable, IC, were examined to determine the overall spread of participants across the measures.

Table 1. IC Dichotomous Rasch Model Summary Statistics

	Ability (θ) (*N*=75)	Item (δ) (*N* = 51)
Calibrations		
Measure (Logits)		
M	−0.03	−1.46
SD	0.91	0.50
Standard Error		
M	0.39	0.31
SD	0.09	0.03
Model-Data Fit		
Infit *MSE*		
M	0.96	0.99
SD	0.09	0.07
Std. Infit		
M	−0.13	0.00
SD	0.67	0.48
Outfit *MSE*		
M	0.98	0.03
SD	0.18	0.18
Std. Outfit		
M	−0.09	0.12
SD	0.84	0.88

Table 1 presents a summary of the results from the analysis of the IC assessments. Specifically, the calibration of participants' ability (e.g., skill with inhibitory control; $N = 75$) and item difficulty (e.g., inhibition required for a trial; $N = 51$) are summarized using average logit-scale calibrations, standard errors, and model-data fit statistics. Examination of the results indicates that, on average, the participants' abilities were located higher on the logit scale ($M = -0.04$, $SD = 1.20$), compared to items ($M = -1.08$, $SD = 0.37$). This finding suggests that the items were relatively easy for the sample of kids who participated in the present study.

As expected, average values of the Standard Error (*SE*) are slightly higher for participants ($M = 0.61$) than items ($M = 0.28$), which may reflect potentially poor targeting for some participants. Average values of model-data fit statistics indicate overall moderate fit to the model and should be near 1.00, indicating little distortion. Values over 1.00 indicate data that is overfit (e.g., over-predictable) while values less than 1.00 indicate a lack of distortion. Specifically, average Infit (inlier-sensitive fit; these values are more sensitive to person response patterns) (Items: $M = 1.00$, $SD = 0.12$; Participants: $M = 0.90$, $SD = 0.18$) and Outfit (Items: $M = 0.90$, $SD = 0.18$); Participants: $M = 0.89$, $SD = 0.20$) mean square statistics were around the expected value of 1.00. Likewise, average standardized Infit

https://doi.org/10.1017/S0305000922000678 Published online by Cambridge University Press

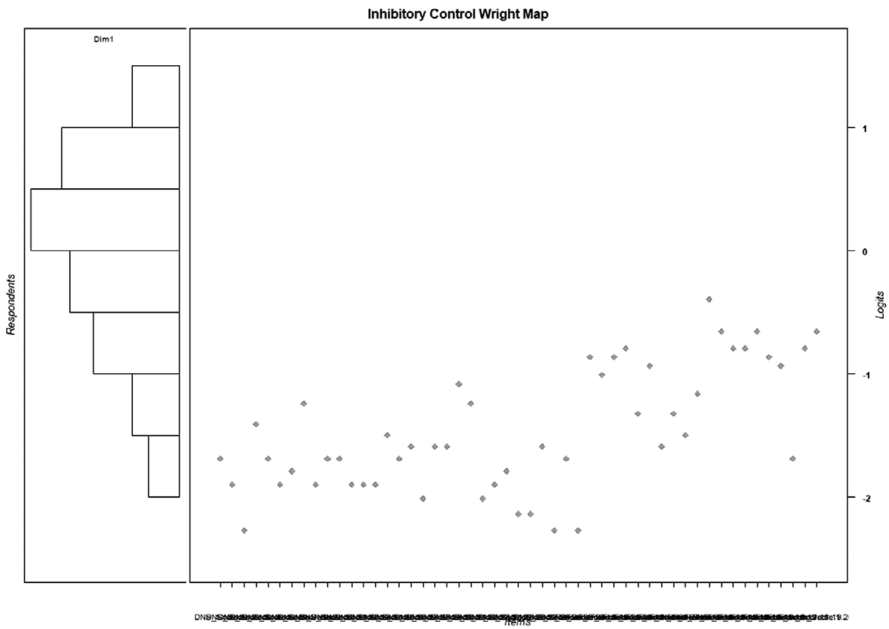

Figure 4. IC Measures Dichotomous Rasch Model Participant and Item Calibration

(Items: $M = -0.01$, $SD = 0.86$; Participants: $M = -0.07$, $SD = 0.45$) and Outfit (e.g., outlier-sensitive fit; these values are more sensitive to item response patterns) (Items: $M = 1.02$, $SD = 0.22$; Participants: $M = 0.89$, $SD = 0.20$) statistics were near the expected value of 0.00 when data fit the model. This finding of adequate fit to the model supports the interpretation of item and person calibrations on the logit scale as indicators of their locations on the latent variable IC.

Figure 4 illustrates the calibrations of the Participants and Items on the logit scale that represents the latent variable. The calibrations shown in this figure correspond to the calibrations for items and persons. The rightmost column (Measure) shows the logit scale. Higher numbers correspond to higher levels of achievement (for participants) and higher levels of difficulty (for items), and lower numbers correspond to lower achievement and less difficulty, respectively, for participants and items.

Next, participants on the latent variable are illustrated using the histogram. Examination of the histogram indicates a wide spread of achievement levels, with most students grouped near the right of the logit scale ($\theta > 0.00$). Next, item locations on the logit scale are plotted on the right side. Examination of the item plotting indicates a similar overall spread as the participants measures. However, the items appear somewhat clustered at the lower half of the logit scale, without many items appearing above the average person location ($\theta >= 0.00$). This lack of moderate-difficulty items may have contributed to the somewhat large SE values for students with middle-range calibrations.

Coding SA task

SA Task responses were coded as "Correct," "Object Error," or "Destination Error." Consider the prompt, "Put the frog on the napkin in the box," illustrated in Figure 1. In the

https://doi.org/10.1017/S0305000922000678 Published online by Cambridge University Press

1-referent condition, the correct response is to place the frog in the box. In the 2-referent condition, the correct response is to place the frog *that's* on the napkin in the box. Any selection of other objects to move (e.g., the empty frog in the 2-referent condition, the napkin, etc.) was coded as an Object Error, as it involved the participant selecting the incorrect object. Any movement of any object to a destination other than the box was coded as a Destination Error (e.g., moving the frog to the napkin or to be with the other frog).

Analysis of correct responses

Overall, children were accurate on roughly 54% of trials ($M = 14.98/28$). For the 14 trials on the 1-referent task, children responded correctly on an average of 8.87 trials. For the 14 trials on the 2-referent task, children responded correctly on an average of 6.12 trials. In order to determine whether the average number of correct trials differed by task (1-referent vs 2-referent) or by IC, a repeated measures ANCOVA was conducted with task as a categorical, within-subjects repeating variable and IC as a continuous covariate. Age did not significantly correlate with percentage correct ($r = -.05$, $p = .68$) and was therefore not included in the model. This analysis found main effects for task, $F(1, 73) = 64.97$, $p < .001$, $\eta_p^2 = .471$, and IC level, $F(1, 73) = 7.93$, $p = .01$, $\eta_p^2 = .10$. There was also a significant two-way interaction between task and IC: $F(1, 73) = 4.93$, $p = .03$, $\eta_p^2 = .06$.

The two-way interaction between task and IC level indicates differences in the importance of IC to 1-referent and 2-referent trials. Although IC significantly correlated with percent correct on 2-referent trials ($r = .36$, $p < .001$), it was marginally significant with percent correct on 1-referent trials ($r = .22$, $p = .06$). The main effect of task indicates that, when IC is controlled for, children made significantly more correct responses on the 1- ($M = .62$, $SD = .03$) than on the 2-referent trials ($M = .43$, $SD = .04$). Overall, the significance of IC as a covariate demonstrates that IC is an important predictor of overall accuracy on SA tasks.

Analysis of errors

Although children were correct more often than they were incorrect, they still made errors on roughly 46% of trials. Next, we explore two common types of errors that lead to incorrect responses: destination and object errors. Age did not significantly correlate with overall percentage incorrect ($r =.05$, $p = .68$), or the percentage of destination ($r = -.002$, $p = .99$) or object errors ($r = .14$, $p = .23$), and was therefore not included in the model.

Destination errors

In order to determine whether the average number of destination errors differed by task, age, or by IC level, a mixed factor repeated measures ANCOVA was conducted with task as a categorical, within-subjects repeating variable and IC was a continuous covariate. This analysis found a main effect of IC, $F(1, 73) = 12.40$, $p < .001$, $\eta^2 = .145$. There was not a main effect of referent, $p = .89$, or a significant interaction between referent and IC, $p = .46$. Finally, the Pearson correlation between a child's IC level and their number of destination errors was significant, $r = -.38$, $p = .001$. The significant main effect of IC indicated that as IC level increased, the percentage of 1-referent and 2-referent errors decreased. This correlation is significant for both 1-referent ($r = -.36$, $p = .002$) and 2-referent ($r = -.39$, $p < .001$) errors.

Object errors

In order to determine whether the average percentage of object errors differed by task or by IC level, a mixed factor repeated ANCOVA was conducted with task as a categorical, within-subjects repeating variable and IC as a continuous covariate. There were significant main effects of IC, $F(1, 73) = 4.51$, $p = .037$, $\eta^2 = .06$, and referent, $F(1, 73) = 64.78$, $p < .001$, $\eta^2 = .47$. However, these main effects were qualified by a significant 2-way interaction between task and IC, $F(1, 73) = 5.55$, $p = .02$, $\eta^2 = .07$. Additionally, IC significantly correlated to overall percentage of object errors ($r = .91$, $p < .001$). This relationship persisted for 2-referent ($r = -.32$, $p = .01$) but not 1-referent trials ($r = -.10$, $p = .40$).

The 2-way interaction between task and IC revealed a more robust relationship between a child's IC level and their percentage of object errors for 2-referent trials compared to 1-referent trials. The Pearson correlation between a child's IC level and their number of object errors was significant, $r = -.25$, $p = .03$. The significance of this correlation held for the 2-referent task ($r = -.32$, $p = .01$), but not for the 1-referent task ($r = -.10$, $p = .40$). The significant main effect of task was driven by a higher percentage of object errors on 2-referent trials ($M = .54$, $SE = .03$) compared to 1-referent ($M = .30$, $SE = .03$) even when IC was controlled for. Finally, the significant main effect of IC as a covariate indicates that IC is a reliable predictor of children's performance on SA tasks.

Discussion

Resolving SA sentences can be cognitively demanding for children. The present study examined children's resolution of 1- and 2-referent SA sentences with accompanying visual scenes. Given the nature of these tasks, participants who failed to correctly resolve a SA sentence could commit an object error, meaning they move the wrong object, or a destination error, meaning they move an object to the wrong destination, or both. Destination errors can occur if children misinterpret the SA sentence and place the correct object (e.g., frog sitting on the napkin) in the wrong destination (e.g., the empty napkin, next to the other frog). Object errors can occur if children misinterpret the SA sentence and move the wrong object (e.g., frog not on the napkin, empty napkin).

Here, we were particularly interested in examining whether a tendency towards certain types of errors might be related to a child's level of IC. We expected participants with higher IC to outperform participants with lower IC because they would be better able to suppress and revise their initial interpretation. We also expected the 1-referent task, in general, to yield better performance than the 2-referent task given the simplified visual scene. Finally, we hypothesized that the types of errors committed might differ by IC level, with lower IC being particularly vulnerable to destination errors in the 1-referent condition and to object errors in the 2-referent condition, reflecting a tendency to perseverate on their first interpretation.

In this study we find some support for each of these hypotheses and thus provide evidence that IC is important in children's resolution of SA sentences. First, our study demonstrates that children performed better on the less demanding 1-referent task compared to the more demanding 2-referent task. This supports existing literature demonstrating that 2-referent SA prompts produce more ambiguity than 1-referent prompts and are therefore more difficult (Altmann, Garnham, & Dennis, 1992; Spivey et al., 2002; Tanenhaus et al., 1995).

Second, our study shows that children with higher IC correctly resolved more SA sentences than children with lower levels of IC. This pattern persisted in both the 1- and

2-referent tasks. This is in line with existing research indicating that children struggle with SA, typically maintaining their early interpretation (Anderson et al., 2011; Choi & Trueswell, 2010; Hurewitz et al., 2000; Kidd & Bavin, 2005; Trueswell & Gleitman, 2004; Trueswell et al., 1999; Weighall, 2008). Still, though, it was common for children to make errors on the SA tasks. In fact, children correctly resolved SA sentences on only 54% of trials. This was expected, based on rates documented in existing literature (54% – Choi & Trueswell, 2010; 60% – Trueswell et al., 1999). Further, even children with high IC averaged errors on at least 8 trials (out of 28 trials, 29% error rate). Our final set of analyses then explored the types of errors children commit, particularly those errors that might be related to IC.

Destination errors

For destination errors, high IC corresponded to a lower overall error percentage. However, we do not find a high frequency of destination errors (M = 5.84/28, or on approximately 20.86% of trials). As a rule, these were the least common errors children made. Our study demonstrates that increases in IC correspond to decreases in the number of destination errors on both 1- and 2-referent trials. Trueswell et al. (1999) argued that a destination bias reflected a difficulty with integrating the verbal prompt with the visual scene. This failure to integrate multiple sources of information may be reflective of children's still developing IC. IC has been shown to facilitate the control of multiple sources of information and maintaining task-relevant information in working memory (Lustig, Hasher, & May, 2001; Munakata, Herd, Chatham, Depue, Banich, & O'Reilly, 2011). Thus, it is plausible that IC facilitates the integration of visual and verbal information on SA tasks, contributing to a child's ability to successfully resolve SA.

Alternatively, IC may assist a child in overriding their tendency to perseverate on their first interpretation of the sentence in order to adapt to a new interpretation. It is well-documented that the tendency to make perseverative errors decreases as IC develops (Gunning-Dixon & Raz, 2003; Stahl & Pry, 2005). Children's tendency towards destination errors may reflect an overall tendency to make perseverative errors in which they cannot override their initial interpretation. Although it is difficult to isolate individual components of IC, an important next step may be to employ a more robust IC battery in a sample with a larger age range in order to better examine how IC contributes to SA resolution.

Object errors

For object errors, higher IC corresponded to a lower overall error rate. Children with higher IC made fewer object errors than children with lower IC. In our study, children's pattern of object errors differed based on the number of referents in the task. For the 1-referent task, the object error rate was similar regardless of a child's IC level. But, for the 2-referent task, children with higher IC had a much lower error rate. Overall, for object errors, only on the more demanding 2-referent task did a child's level of IC differentiate error rates. The 1-referent condition produces ambiguity after hearing "on the napkin," as it is appropriate to say "the frog" because there is only one frog. Of course, in the 1-referent condition there is just a single target object, which makes an object error unlikely. In the 1-referent condition, if the child were to stop at the initial interpretation, they would put the frog on the napkin. This would result in a destination error because

https://doi.org/10.1017/S0305000922000678 Published online by Cambridge University Press

they did not listen to the remaining prompt and acted preemptively, placing the frog on the napkin rather than on the box, which would be the correct interpretation. Conversely, the 2-referent condition produces ambiguity in identifying which frog should be moved. Thus, although destination errors likewise increase, object errors are significantly more frequent in the 2-referent condition. A failure to revise the initial interpretation in the 2-referent condition might result in the child moving the lone frog rather than the frog sitting on the napkin (an object error). In line with existing literature, this demonstrates that 2-referent SA prompts produce more ambiguity than 1-referent prompts and are therefore more difficult (Altmann et al., 1992; Spivey et al., 2002; Tanenhaus et al., 1995). The present study extends this, demonstrating that the number of referents likewise affects the type of error a child is more likely to commit.

Though not directly assessed in the present study, a high number of object errors may reflect poor WM. An object error would occur when a child is unable to maintain the beginning of the sentence in memory as incoming information is processed. In essence, children might process the sentence as if they heard, “Put the *blah blah blah* in the box.” For the 1-referent task, this could lead children to select the wrong object (the napkin or distractor) but move it to the right destination (the box). Similarly, for the 2-referent task children could also move the wrong object (the lone frog or the napkin in this case) to the right destination. Of the individual components that make up EF, the tasks that we use are argued to predominantly implicate IC. Although it is challenging to determine the precise EF component that may be most at play here because the task is not pure, and it is difficult to isolate individual components of EF, a next step may be to determine whether WM specifically is driving the emerging pattern evidenced in the present paper.

Alternatively, it may be that the high number of object errors in children with lower IC points to a lack of experience using IC within language specifically. Recently, Doebel (2020) argued that IC instead is the development of skills in using control towards specific goals that are embedded in a particular knowledge set, culture, sociocultural context, belief, or group. Similarly, the engagement and motivational significance to an individual likewise affects a child’s ability to engage control towards a specific goal. As such, it may be that the high number of object errors in children with lower IC may not reflect lower IC per se, but a lack of experience in using IC in manipulating language. Both receptive and expressive language skills during preschool predict later EF levels, even when socio-economic and demographic factors are controlled for (Fuhs & Day, 2011; Schmitt, Purpura, & Elicker, 2019). Moreover, bidirectional effects between language and EF are extensively documented (Gandolfi & Viterbori, 2020; Gooch, Thompson, Nash, Snowling, & Hulme, 2016). It may be that as children acquire language experience, they improve their ability to revise their initial response such that it is language and not EF per se that impact performance on SA tasks. Future work may benefit from further exploring the role of IC and EF in SA through both verbal and non-verbal EF tasks to compare how performance on a diverse, comprehensive, and ecologically valid EF battery relates to SA performance.

Conclusions

Overall, the present study suggests that successful SA resolution relies not only on the development of linguistic skills, such as parsing principles, but on domain-general cognitive skills like IC. This extends existing studies demonstrating the importance of

IC for SA (Caplan, Alpert, Waters, & Olivieri, 2000; Choi & Trueswell, 2010; Novick et al., 2005; Trueswell et al., 2011; Weighall, 2008; Woodard et al., 2016) and supports research citing a role for IC in a variety of language processes (Bishop, Nation, & Patterson, 2014; Fuhs et al., 2014; Gooch et al., 2016; Karasinski, 2015; Slot & von Suchodoletz, 2018). The current study adds to this body of work supporting a need for IC in language processes, demonstrating a role for IC in SA resolution. To the best of our knowledge, no other research to date has examined IC in relation to specific SA error types. Thus, additional research examining this effect is needed.

In sum, this study provides three lines of evidence supporting the role of IC in SA resolution. First, we show that children with higher IC resolve more SA sentences correctly. Second, we show that SA resolution is worse on tasks that place higher demands on IC, even for children with high IC. Third, we show that children with higher IC make different types of SA errors than children with lower IC. In sum, this study expands theoretical understanding of SA resolution, demonstrating that IC is important not only to SA accuracy but to the types of errors children are prone to.

Supplementary Materials. To view supplementary material for this article, please visit http://doi.org/10.1017/S0305000922000678.

References

Altmann, G. T., **Garnham, A.**, & **Dennis, Y.** (1992). Avoiding the garden path: Eye movements in context. *Journal of Memory and Language*, **31**(5), 685–712. doi:10.1016/0749-596x(92)90035-v

Anderson, S. E., **Farmer, T. A.**, **Goldstein, M.**, **Schwade, J.**, & **Spivey, M. J.** (2011). Individual differences in measures of linguistic experience account for variability in the sentence processing skill of five-year-olds. *Experience, Variation and Generalization Trends in Language Acquisition Research,* 203-222. https://doi.org/10.1075/tilar.7.12and

Bishop, D. V., **Nation, K.**, & **Patterson, K.** (2014). When words fail us: insights into language processing from developmental and acquired disorders. *Philosophical transactions of the Royal Society of London. Series B, Biological Sciences*, **369**(1634), 20120403. https://doi.org/10.1098/rstb.2012.0403

Caplan, D., **Alpert, N.**, **Waters, G.**, & **Olivieri, A.** (2000). Activation of Broca's area by syntactic processing under conditions of concurrent articulation. *Human Brain Mapping*, **9**(2), 65–71. https://doi.org/10.1002/(SICI)1097-0193(200002)9:2<65::AID-HBM1>3.0.CO;2-4

Choi, Y., & **Trueswell, J. C.** (2010). Children's (in)ability to recover from garden paths in a verb-final language: Evidence for developing control in sentence processing. *Journal of Experimental Child Psychology*, **106**(1), 41–61. doi:10.1016/j.jecp.2010.01.003

Christianson, K., **Hollingworth, A.**, **Halliwell, J. F.**, & **Ferreira, F.** (2001). Thematic Roles Assigned along the Garden Path Linger. *Cognitive Psychology*, **42**(4), 368–407. doi:10.1006/cogp.2001.0752

Diamond, A. (2013). Executive functions. *Annual Review of Psychology*, **64**, 135–168. doi:10.1093/acprof:oso/9780195169539.003.0006

Doebel, S. (2020). Rethinking executive function and its development. *Perspectives on Psychological Science*, **15**(4), 942–956.

Eberhard, K. M., **Spivey-Knowlton, M. J.**, **Sedivy, J. C.**, & **Tanenhaus, M. K.** (1995). Eye movements as a window into real-time spoken language comprehension in natural contexts. *Journal of Psycholinguistic Research*, **24**(6), 409–436. doi:10.1007/BF02143160

Engelhardt, P. E., **Nigg, J. T.**, & **Ferreira, F.** (2017). Executive Function and Intelligence in the Resolution of Temporary Syntactic Ambiguity: An Individual Differences Investigation. *Quarterly Journal of Experimental Psychology*, **70**(7), 1263–1281. doi:10.1080/17470218.2016.1178785

Fuhs, M. W., & **Day, J. D.** (2011). Verbal ability and executive functioning development in preschoolers at head start. *Developmental psychology*, **47**(2), 404. https://doi.org/10.1037/a0021065

Fuhs, M. W., **Nesbitt, K. T.**, **Farran, D. C.** & **Dong, N.** (2014). Longitudinal associations between executive functioning and academic skills across content areas. *Developmental Psychology*, **50**(6), 1698–1709. https://doi.org/10.1037/a0036633

Gandolfi, E., & **Viterbori, P.** (2020). Inhibitory control skills and language acquisition in toddlers and preschool children. *Language Learning*, **70**(3), 604–642. 10.1111/lang.12388

Gerstadt, C. L., **Hong, Y. J.**, & **Diamond, A.** (1994). The relationship between cognition and action: performance of children 312–7 years old on a stroop-like day-night test. *Cognition*, **53**(2), 129–153. doi:10.1016/0010-0277(94)90068-x

Gooch, D., **Thompson, P.**, **Nash, H. M.**, **Snowling, M. J.**, & **Hulme, C.** (2016). The development of executive function and language skills in the early school years. *Journal of Child Psychology and Psychiatry*, **57**(2), 180–187. doi:10.1111/jcpp.12458

Gunning-Dixon, F. M., & **Raz, N.** (2003). Neuroanatomical correlates of selected executive functions in middle-aged and older adults: a prospective MRI study. *Neuropsychologia*, **41**(14), 1929–1941. https://doi.org/10.1016/S0028-3932(03)00129-5

Hsu, N. S., & **Novick, J. M.** (2016). Dynamic engagement of cognitive control modulates recovery from misinterpretation during real-time language processing. *Psychological Science*, **27**(4), 572–582. doi: 10.1177/0956797615625223

Hurewitz, F., **Brown-Schmidt, S.**, **Thorpe, K.**, **Gleitman, L. R.**, & **Trueswell, J. C.** (2000). One frog, two frog, red frog, blue frog: Factors affecting children's syntactic choices in production and comprehension. *Journal of Psycholinguistic Research*, **29**(6), 597–626. doi:10.1023/A:1026468209238

Hussey, E. K., & **Novick, J. M.** (2012). The Benefits of Executive Control Training and the Implications for Language Processing. *Frontiers in Psychology*, **3**. doi:10.3389/fpsyg.2012.00158

Karasinski, C. (2015). Language ability, executive functioning and behavior in school-age children. *International Journal of Language & Communication Disorders*, **50**(2), 144–150. doi:10.1111/1460-6984.12104

Kidd, E., & **Bavin, E. L.** (2005). Lexical and referential cues to sentence interpretation: An investigation of children's interpretations of ambiguous sentences. *Journal of Child Language*, **32**(4), 855–876. doi: 10.1017/s0305000905007051

Kidd, E., & **Holler, J.** (2009). Children's use of gesture to resolve lexical ambiguity. *Developmental Science*, **12**(6), 903–913. https://doi.org/10.1111/j.1467-7687.2009.00830.x

Kim, J. H., & **Christianson, K.** (2013). Sentence complexity and working memory effects in ambiguity resolution. *Journal of Psycholinguistic Research*, **42**(5), 393–411. doi:10.1007/s10936-012-9224-4

Knight, R. T., & **Stuss, D. T.** (2002). Prefrontal cortex: The present and the future. *Principles of Frontal Lobe Function*, 573–597. doi:10.1093/acprof:oso/9780195134971.003.0034

Livesey, D., **Keen, J.**, **Rouse, J.**, & **White, F.** (2006). The relationship between measures of executive function, motor performance and externalizing behavior in 5-and 6-year-old children. *Human Movement Science*, **25**(1), 50–64. doi:10.1016/j.humov.2005.10.008

Lustig, C., **Hasher, L.**, & **May, C.** (2001). Working memory span and the role of proactive interference. *Journal of Experimental Psychology: General*, **130**, 199–207. doi:10.1037/0096-3445.130.2.199

Macdonald, M. C., **Pearlmutter, N. J.**, & **Seidenberg, M. S.** (1994). The lexical nature of syntactic ambiguity resolution. *Psychological Review*, **101**(4), 676–703. doi:10.1037/0033-295x.101.4.676

Meroni, L., & **Crain, S.** (2003). On not being led down the kindergarten path. In *Proceedings of the 27th Boston University Conference on Language Development* (pp. 531–544). Somerville, MA: Cascadilla Press.

Munakata, Y., **Herd, S. A.**, **Chatham, C. H.**, **Depue, B. E.**, **Banich, M. T.**, & **O'Reilly, R. C.** (2011). A unified framework for inhibitory control. *Trends in cognitive sciences*, **15**(10), 453–459. doi:10.1016/j.tics.2011.07.011

Novick, J. M., **Trueswell, J. C.**, & **Thompson-Schill, S. L.** (2005). Cognitive control and parsing: Reexamining the role of Broca's area in sentence comprehension. *Cognitive, Affective, & Behavioral Neuroscience*, **5** (3), 263–281. doi:10.3758/cabn.5.3.263

Novick, J. M., **Hussey, E.**, **Teubner-Rhodes, S.**, **Harbison, J. I.**, & **Bunting, M. F.** (2013). Clearing the garden-path: Improving sentence processing through cognitive control training. *Language, Cognition and Neuroscience*, **29**(2), 186–217. doi:10.1080/01690965.2012.758297

Qi, Z., **Love, J.**, **Fisher, C.**, & **Brown-Schmidt, S.** (2020). Referential context and executive functioning influence children's resolution of syntactic ambiguity. *Learning, Memory, and Cognition*, **46**(10), 1922–1947. https://doi.org/10.1037/xlm0000886

Rasch, G. (1960). *Studies in mathematical psychology: I. Probabilistic models for some intelligence and attainment tests*. Oxford, England: Nielsen & Lydiche.

Robitzsch, A., & **Lüdtke, O.** (2020). A review of different scaling approaches under full invariance, partial invariance, and noninvariance for cross-sectional country comparisons in large-scale assessments. *Psychological Test and Assessment Modeling*, **62**(2), 233–279.

Schmitt, S. A., **Purpura, D. J.**, & **Elicker, J. G.** (2019). Predictive links among vocabulary, mathematical language, and executive functioning in preschoolers. *Journal of Experimental Child Psychology*, **180**, 55–68. https://doi.org/10.1016/j.jecp.2018.12.005

Slot, P. L., & **von Suchodoletz, A.** (2018). Bidirectionality in preschool children's executive functions and language skills: Is one developing skill the better predictor of the other? *Early Childhood Research Quarterly*, **42**, 205–214. doi:10.1016/j.ecresq.2017.10.005

Snedeker, J., & **Trueswell, J. C.** (2004). The developing constraints on parsing decisions: the role of lexical-biases and referential scenes in child and adult sentence processing. *Cognitive Psychology*, **49**(3), 238–299. https://doi.org/10.1016/j.cogpsych.2004.03.001

Spivey, M., **Tanenhaus, M. K.**, **Eberhard, K. M.**, & **Sedivy, J. C.** (2002). Eye movements and spoken language comprehension: Effects of visual context on syntactic ambiguity resolution. *Cognitive Psychology*, **45**(4), 447–481. doi:10.1016/s0010-0285(02)00503-0

Stahl, L., & **Pry, R.** (2005). Attentional flexibility and perseveration: Developmental aspects in young children. *Child Neuropsychology*, **11**(2), 175–189. https://doi.org/10.1080/092970490911315

Tanenhaus, M. K., **Spivey-Knowlton, M. J.**, **Eberhard, K. M.**, & **Sedivy, J. C.** (1995). Integration of visual and linguistic information in spoken language comprehension. *Science*, **268**(5217), 1632–1634. doi: 10.1126/science.7777863

Traxler, M. J., **Morris, R. K.**, & **Seely, R. E.** (2002). Processing subject and object relative clauses: Evidence from eye movements. *Journal of Memory and Language*, **47**(1), 69–90. https://doi.org/10.1006/jmla.2001.2836

Trueswell, J., & **Gleitman, L.** (2004). Children's Eye Movements during Listening: Developmental Evidence for a Constraint-Based Theory of Sentence Processing. In J. M. Henderson & F. Ferreira (Eds.), *The Interface of Language, Vision, and Action: Eye Movements and the Visual World* (pp. 319–346). Psychology Press.

Trueswell, J. C., **Papafragou, A.**, & **Choi, Y.** (2011). Referential and syntactic processes: What develops? *The Processing and Acquisition of Reference*, 65–108. doi: 10.7551/mitpress/9780262015127.001.0001

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**(2), 89–134. doi:10.1016/s0010-0277(99)00032-3

Weighall, A. R. (2008). The kindergarten path effect revisited: Children's use of context in processing structural ambiguities. *Journal of Experimental Child Psychology*, **99**(2), 75–95. doi:10.1016/j.jecp.2007.10.004

Wimmer, M. C., **Doherty, M. J.**, & **Collins, W. A.** (2011). The development of Ambiguous Figure Perception. *Monographs of the Society for Research in Child Development*, **76**(1), i–130. http://www.jstor.org/stable/41261555

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. doi:10.1016/j.jecp.2015.08.005

Cite this article: May, K.E., & **Scofield, J.** (2024). Inhibitory Control and Patterns of Errors in Resolution of Syntactically Ambiguous Sentences. *Journal of Child Language* **51,** 271–287, https://doi.org/10.1017/S0305000922000678