

The extragrammaticality of the acquisition of adjunct control

Juliana Gerard

School of Communication and Media, Ulster University, Jordanstown, UK

Previous research on 4-6-year olds' interpretations of adjunct control has observed non-adultlike behavior for sentences like *John called Mary before running to the store*. Several studies have aimed to identify a grammatical source of children's errors. This study tests the predictions of grammatical and extragrammatical accounts by comparing children's behavior on two truth value judgment tasks: a high demand task, with a true/false judgment based on event ordering; and a low demand task, with a true/false judgment based on the color of an item. Children's behavior is more adultlike on the low demand task, suggesting that children's interpretations may be influenced by extragrammatical factors. Implications are discussed for children's behavior in previous studies and for the role of the linguistic input.

Keywords: adjunct control; task effects; TVJT

Introduction

Children's interpretations are sometimes different from adults' interpretations. Sometimes the differences are minor, with no significant disruption in communication; sometimes children's interpretations deviate quite a bit, and impact their understanding in a conversation. Importantly, any differences are generally not random: children's non-adultlike interpretations are a sign of some particular difference between children and adults in generating an interpretation of a sentence. Depending on what that difference is, there are different expectations for children's interpretations in the moment, but also for acquisition.

This paper considers these expectations in the context of adjunct control, as in (1).

(1) John called Mary before PRO running to the store.

In (1), the silent adjunct subject, notated here as PRO, is bound by the main clause subject John but not the object Mary, and adults typically only accept a subject control interpretation (that John ran to the store). Meanwhile, children in previous studies have accepted a wider range of interpretations.

One option for children's non-adultlike interpretations is a non-adultlike grammar, which generates a superset of the interpretations allowed by the adult grammar (Goodluck 1981; Hsu, Cairns & Fiengo 1985; Hsu et al. 1989; McDaniel, Cairns & Hsu 1991; Goodluck & Behne 1992; Wexler 1992; Cairns et al. 1994; Broihier & Wexler 1995; Adler 2006; see also Landau 2021). However, another option is that children erroneously retrieve an ungrammatical antecedent due to extragrammatical factors (other than syntactic knowledge) (Gerard et al. 2017; Gerard et al. 2018).

Grammatical and extragrammatical accounts have different implications for children's acquisition, for control and more generally for syntactic dependencies: if children's grammars are not adultlike, then evidence must be available in the input for the adult grammar. However, different evidence is expected in the input under an extragrammatical account. In this paper, the predictions of grammatical and extragrammatical accounts in an experimental context for children's behavior. Based on these predictions, further implications may be considered for the linguistic input.

The following sections discuss the patterns of interpretations observed in studies on the acquisition of adjunct control and proposed accounts. Predictions of grammatical and extragrammatical accounts are tested using a truth value judgment task, and implications are discussed for how control is acquired.

Previous studies

Previous studies on the acquisition of adjunct control have consistently observed non-adultlike behavior for sentences like (1). In particular, children have allowed a wider range of antecedents for adjunct PRO than in the adult grammar (strict subject control).

Across studies, the following non-adult response patterns have been reported for adjunct control (Goodluck 1981; Hsu, Cairns & Fiengo 1985; Lust et al. 1986; Hsu et al. 1989; McDaniel, Cairns & Hsu 1991; Goodluck & Behne 1992; Cairns et al. 1994; Broihier & Wexler 1995; Goodluck 2001; Adler 2006; Gerard et al. 2017; Gerard et al. 2018; Janke 2018):

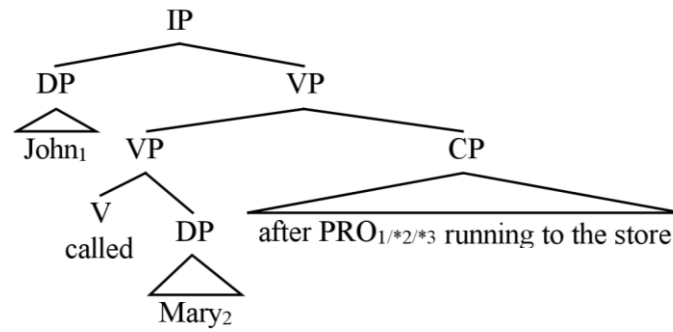
- (2) a. free reference: control by any referent in the task context
- b. object control: strict control by the main clause object (*Mary* in (1))
- c. control by an internal antecedent, but not an external antecedent¹¹

In general, these non-adultlike patterns of behavior have been attributed to non-adultlike grammars, with one or more non-adultlike stages before the adult grammar.

In the adult grammar, PRO is controlled by the closest c-commanding NP. Along with the attachment height of the adjunct, this identifies the main clause subject in sentences like (1) as the antecedent of adjunct PRO:

¹ This pattern of behavior has been observed in adults and older children in contexts with pragmatic pressure for the object as a discourse topic (Janke & Bailey, 2017; Janke & Perovic, 2017; Janke, 2018); meanwhile, younger children who accept object control do so even in a neutral discourse context.

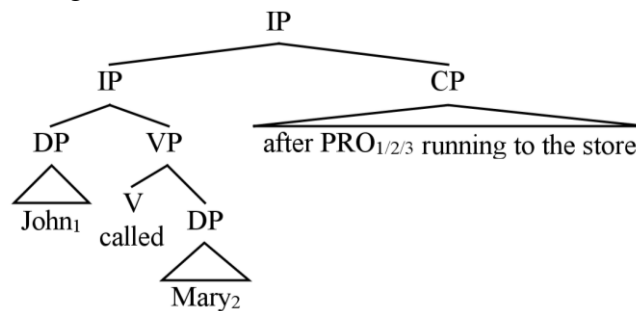
(3)



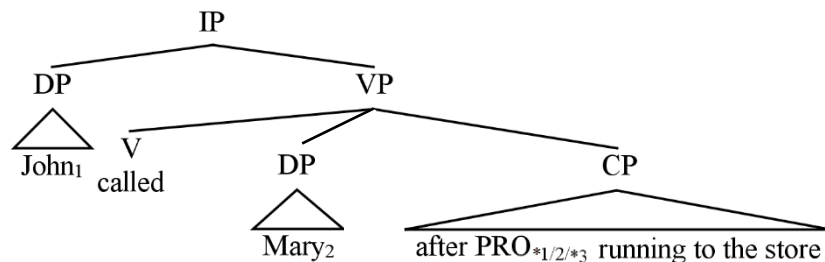
Children's non-adult behavior may therefore be accounted for by a grammar that lacks these components.

For example, under the Variable Attachment account (Goodluck 1981; Hsu, Cairns & Fiengo 1985; McDaniel, Cairns & Hsu 1991; Cairns et al. 1994; Adler 2006), children misattach the adjunct to the main clause, as in (4):

(4) a. High attachment:



b. Low attachment:



In the adult grammar, the main clause subject binds PRO, while the object does not. In contrast, the high attachment in (4a) disallows binding by any internal arguments, requiring an antecedent for PRO to be retrieved from the discourse. Meanwhile, the low attachment in (4b) allows the main clause object to bind PRO (i.e. object control). As

PRO is the *closest* c-commanding antecedent, low attachment also rules out subject control. These attachments predict the distinct patterns of behavior in (2): high attachment predicts free reference; low attachment predicts strict object control; and optional low attachment predicts control by internal referents (*John* and *Mary* in (1)), but not by external referents.

These patterns have all been reported, although at various rates across studies. Additionally, with a grammar that allows free reference (a *free reference* grammar), even if any referent is a possible antecedent, the limited number of observations in a given task may not include all possible antecedents (Goodluck & Behne 1992; Wexler 1992; McDaniel & Cairns 1990a). Therefore, the strict object and sentence-internal patterns in (2) may also be accounted for by a single non-adult grammar which allows free-reference, with children's interpretations for a given context determined by discourse factors (Adler 2006; Wexler 1992; Broihier & Wexler 1995; Goodluck 2001; Goodluck & Behne 1992; Goodluck 1998).

In addition to high attachment (Adler 2006), a pattern of free reference is predicted if children's grammars lack a syntactic restriction on the controller (Goodluck & Behne 1992), or if they lack PRO, resulting in a nominal structure as in (5):

(5) John called Mary before [_{NP} (the) running to the store]

In (5), the agent of *running* is underspecified in the adult grammar. This predicts the same free reference pattern of behavior as a grammar with high attachment or which does not restrict the controller to the closest c-commanding NP. Thus, both high attachment and a nominal structure are consistent with children's behavior in previous studies; however, the variation across studies remains unaccounted for (Table 1).

Table 1: Previous studies on the acquisition of adjunct control.

Task	Study	substudy	subject %	adultlike %	n adultlike	n	age range
act out	Hsu et al (1985)			23%	15	64	3;2-8;3
		4;1-4;6	14%			7	4;1-4;6
		4;7-5;0	38%			7	4;7-5;0
		5;1-5;6	46%			8	5;1-5;6
		5;7-6;0	29%			7	5;7-6;0
	Lust et al 1986	no lead	27%			101	3;1-7;11 (m = 5;5)
		pragmatic lead	18%			101	3;1-7;11 (m = 5;5)
	Hsu et al (1989)			26%	21	81	4;7-8;0
	Goodluck 1981	study 1: 4yo	45%			10	4
		study 1: 5yo	60%			10	5
		study 1: 6yo	67%			10	6
		study 2	65%			20	4-5
	Goodluck (1998)		90%			24	4-5
	Goodluck (2001)	study 1	28%	15%	3	20	4;5-6;2 (m = 5;7)
		study 2	71%	65%	13	20	4;4-6;5 (m = 5;8)
	Goodluck & Behne (1992)	4 year olds	89%			13	4
		5 year olds	92%			16	5
		6 year olds	81%			13	6
grammaticality judgment	McDaniel et al 1991	study 1 (act out)		10%	2	20	4;0-5;1
		study 1 (judgment)		30%	6	20	4;0-5;1
		time 1		29%	4	14	4;1-4;9
		time 2		29%	4	14	4;4-5;0
		time 3		50%	7	14	4;7-5;4
	Cairns et al (1994)	time 1		0%	0	14	3;10-4;11
		time 2		33%	5	15	4;2-5;3
		time 3		33%	5	15	4;7-5;8
Picture Selection	Janke (2018)	year 2	78%			14	6;9-7;8
		year 3+	92%			15	8;0-8;7
Coloring	Gerard et al (2017)	gender match	60%	50%	12	24	3;11-5;3 (m = 4;9)
		gender mismatch	73%	71%	17	24	3;11-5;3 (m = 4;9)
		number match	65%	42%	10	24	4;0-5;5 (m = 4;9)
		number mismatch	76%	58%	14	24	4;1-5;5 (m = 4;11)

TVJT	Gerard et al (2018)	Coloring	85%	84%	27	32	4;0-5;3 (m = 4;9)
		TVJT	60%	44%	15	34	4;0-5;3 (m = 4;7)
	Broiheir & Wexler (1995)			43%	6	14	3;10-5;6
	Adler (2006)	study 1	65%			23	3;0-5;11 (m = 4;4)
		study 2	82%			30	3-6 (m = 4;6)

Children's interpretations have been reported based on the overall proportion of subject interpretations in the sample, but also based on the proportion of subject interpretations in each child. Children are categorized by these proportions as having a non-adult grammar (which may include multiple non-adult stages) or the adult grammar.

Typically, children who give at least 75% subject interpretations are categorized as adultlike for adjunct control (Hsu, Cairns & Fiengo 1985; Hsu et al. 1989; McDaniel, Cairns & Hsu 1991; Cairns et al. 1994).²

Variation is observed in both measures across studies. One possible source of this variation is the distribution of children's grammars, if different samples had different proportions of adult grammars. However, more variation is observed across studies than would be expected in the distribution of grammars alone (Conroy et al. 2009).

Another possibility is that the variation is due to extragrammatical factors; that is, any factors aside from syntactic control. For example, for the non-adult grammars which predict a free reference pattern of behavior (or *free reference* grammars),

² This cutoff is acknowledged by Hsu et al (Hsu et al., 1985) to be arbitrary, but a preferred alternative to requiring 100% subject interpretations due to the inevitability of performance errors in children's responses, even with the adult grammar (p. 36).

children's interpretations in any single context will depend on discourse factors, as well as other recent structures in the discourse which may influence the antecedent (Goodluck & Behne 1992; Goodluck 1998; Goodluck 2001). Variation in these factors across tasks will result in different patterns of responses.

Meanwhile, extragrammatical factors may also affect adults' interpretations of adjunct control (Parker, Lago & Phillips 2015; Kush & Dillon 2020), and ceiling accuracy is rarely expected in studies with children, especially for sentences with complex structures, like adjunct control (Omaki & Lidz 2015). Therefore, a further option is that children's grammars are adultlike, and children's non-adultlike behavior for adjunct control is also due to extragrammatical factors.

In previous studies, children's non-adultlike behavior has been attributed non-adult grammars; these grammars are consistent with children's behavior, but grammatical factors alone do not predict the variation across studies. The goal of the current study is to tease apart the roles of grammatical and extragrammatical factors in the acquisition of adjunct control. The following sections discuss the influence of extragrammatical factors for adjunct control across studies. Next, to address the source of children's non-adultlike behavior, the predictions of grammatical and extragrammatical factors are tested in an experimental manipulation of specific task demands. The results are consistent with an extragrammatical account of children's interpretations, but not with a grammatical account. Implications are discussed for the variation in previous studies, and for the role of extragrammatical factors in acquisition from the linguistic input.

Extragrammatical factors

Extragrammatical factors play a role in parsing and producing a response to a linguistic stimulus, and are external to the grammar. This encompasses a range of factors, which

may be categorized based on how and when they influence behavior (Crain & Thornton 1998). For example, in an experimental context, these factors include the parsing processes involved in generating an interpretation of a test sentence, pragmatics and discourse considerations for the test sentence in the task context, and the specific task demands for generating a behavioral response (Crain & Thornton 1998). The following sections consider these factors in the context of adjunct control, as a potential source of variation across studies.

Parsing factors

As the test sentence is delivered, parsing procedures are deployed to generate an interpretation (Frazier & Fodor 1978; Lewis & Vasishth 2005). However, various factors may affect the deployment of these processes, independent of the grammar. For example, sentences with adjunct control involve encoding, storing, and retrieving the antecedent of PRO; meanwhile, the memory mechanisms involved in these operations are affected by factors that have varied across studies.

In (1), repeated below as (6), the verb *running* occurs without an overt subject. This triggers memory mechanisms to search for an antecedent (Lewis & Vasishth 2005; Engelmann, Jäger & Vasishth 2019).

(6) John called Mary before PRO running to the store.

This search is constrained based on syntactic features of the sentence (Dillon et al. 2013; Jäger, Engelmann & Vasishth 2017; Kush & Dillon 2020); however, it is also susceptible to interference from similar referents, or *similarity-based interference* (Lewis & Vasishth 2005; Jäger et al. 2015; Laurinavichyute et al. 2017; Villata, Tabor & Franck 2018; for reviews see Jäger, Engelmann & Vasishth 2017; Jäger et al. 2020; Engelmann, Jäger & Vasishth 2019).

In previous studies on adjunct control, the main clause subject has been similar to other non-subject referents on a number of features (where *similarity* is defined in terms of overlapping features).³ For example, the referents have generally matched in number (singular), animacy (animate), and sometimes also in gender, as in (7):

(7) a. John called Bill before PRO running to the store.

b. Jill called Mary before PRO running to the store.

Meanwhile, children exhibit higher accuracy for adjunct control when the subject and object mismatch in number or gender than when they match (Gerard et al. 2017), with parallel effects observed in adults for gender and animacy (Parker, Lago & Phillips 2015).

Similarity between referents has varied across studies. This suggests that some of the variation in response patterns for adjunct control may be accounted for by similarity-based interference.

A key term here is *some*, rather than all of the variation across studies: importantly, children's interpretations may also be affected by other extragrammatical factors, in addition to similarity-based interference. For example, lexical retrieval cost does not affect the grammaticality of control; however, in a given task, the retrieval cost for different referents may vary due to frequency or word length (Gierut 2001; German & Newman 2004; Storkel & Morrisette 2002; Troia, Roth & Yeni-Komshian 2018;

³ This notion of similarity is also observed under the Relativized Minimality framework (Rizzi, 1990; Friedmann et al., 2009; Rizzi, 2018), in contexts where an ungrammatical antecedent intervenes structurally between the dependent elements. For adjunct control, PRO is the closest c-commanding NP, so a sentence-internal intervener must be linear rather than structural. Relativized Minimality is therefore not considered in the current paper for adjunct control, although further research may consider its implications with a direct comparison between structural and linear interveners (e.g. Friedmann & Costa, 2010).

Key, Jones & Peters 2016). Another consideration is the discourse context for a sentence with adjunct control, discussed further in the following section.

Discourse factors

The adult grammar requires strict subject control for sentences with adjunct control, as in (6). However, both children and adults may access a non-subject interpretation depending on the discourse status of the possible antecedents.

In a discourse neutral context, adults generally access a subject control interpretation for adjunct control. However, when the main clause object is a strongly established topic, object control becomes as frequent as subject control (Janke & Bailey 2017; Janke 2018). Therefore, the acquisition of adjunct control involves a discourse component, in addition to a syntactic restriction on PRO:

- (a) with the adult grammar, strict subject control is expected in a neutral context
- (b) with adultlike sensitivity to the discourse, interpretations may be influenced by a discourse topic

Consequently, non-adultlike behavior may be due to a non-adult grammar, or to non-adultlike sensitivity to the discourse (Conroy et al. 2009; Landau 2021).

Adultlike discourse sensitivity. In previous studies on adjunct control, the discourse context is generally intended to be neutral. An important exception is Janke's (2018) manipulation of a pragmatic lead just before the test sentence (as in (8)), which identified a developmental trajectory of adultlike sensitivity to the discourse in older children (6;9-11;8).

(8) a. Weak pragmatic lead:

Let me tell you something about Hermione. Ron kissed Hermione while flying the broom.

b. Strong pragmatic lead:

Hermione is preparing for a competition. Hermione practises in the air.
Ron kissed Hermione while flying the broom.

(adapted from Janke, 2018)

While the youngest children (6-7 year-olds) were not affected by a pragmatic lead, an increasingly adultlike effect of lead was observed over age, with the oldest age groups (9-11 year-olds) showing adultlike sensitivity to the discourse (i.e. subject control with no lead and a weak lead, but flexible interpretations with a strong lead).

Meanwhile, most studies on the acquisition of adjunct control have focused on children aged 6 and younger. Since younger children's interpretations were not affected by the pragmatic lead, the discourse topic (based on a pragmatic lead) is not likely to be a significant factor in the variation across previous studies.⁴

⁴ Sensitivity to the discourse topic has also been tested by Lust et al (1986) in children aged 3;1-7;11, for adjunct control as well as finite adjuncts. In addition, the position of the adjunct was manipulated, with postposed "forward" adjuncts, as in (6), and preposed "backward" adjuncts, as in (i):

(i) Before PRO running to the store, John called Mary.

Although an effect of pragmatic lead is reported overall, Lust et al (1986) observe that "the [pragmatic lead] has its strongest effect in depressing choice of subject and increasing choice of object on the backward forms, rather than on the forward" (p. 265). This is consistent with the age effect observed by Janke (2018) for forward (postposed) adjuncts, i.e. that younger children are not affected by the pragmatic lead.

Non-adultlike discourse sensitivity. Before adultlike sensitivity is developed to a pragmatic lead, younger children's interpretations may still be influenced by other discourse factors. One such factor is the availability of each referent in the discourse (Hamburger & Crain 1984; Crain & Thornton 1998; Conroy et al. 2009). If an ungrammatical antecedent is more available than the grammatical one, then children may access an ungrammatical interpretation, even with the adult grammar.

For example, in the Truth Value Judgment Task (TVJT) a context is set up which allows for multiple possible interpretations, given a test sentence. In the context, one possible interpretation is false, while the other is true. It is assumed that if both interpretations are available in a child's grammar, then the child will access the true interpretation (the Principle of Charity; Crain & Thornton 1998; Davidson 1984; Grice 1975). Therefore, if a child judges the sentence to be false, this is taken as evidence that the true interpretation is not possible in the child's grammar.

The assumption of the Principle of Charity does not always hold, however. If the true interpretation is not possible in the child's grammar but is more salient in the context than the false interpretation, then the test sentence may still be accepted as true, due to the inaccessibility of the false interpretation (Crain & Thornton 1998; Conroy et

Age effects have also been observed for complement control with the verb *promise* as in (ii) (Sherman & Lust, 1993), which has a similar developmental trajectory to adjunct control (Chomsky, 1969; for a review see Martin, 2016):

- (ii) John promised Mary PRO to run to the store.

While the pragmatic lead affected children's interpretations of *promise* sentences in the oldest age group (7;0-7;11), younger children (3;0-3;11 and 5;0-5;11) did not show an effect of pragmatic lead (Sherman & Lust, 1993). This aligns with the results for adjunct control, i.e. that younger children's interpretations are not influenced by a pragmatic lead with postposed adjuncts.

al. 2009). The reverse case may also be observed: if the false interpretation is not possible in the child's grammar but is more salient than the true interpretation, then the test sentence may still be rejected as false, due to the inaccessibility of the true interpretation.

In previous studies on adjunct control, most children have accepted a subject control interpretation (c.f. Hsu, Cairns & Fiengo 1985; Hsu et al. 1989), but more variation is observed in the rate at which children have rejected a non-subject interpretation (generally the main clause object or an external referent). This suggests that some of the variation in previous studies is due to the salience of non-subject referents - in addition to the parsing factors discussed above.

Meanwhile, one source of variation across tasks is the type of behavioral response (act out, judgment, etc.). Thus, while children's responses are influenced by parsing processes and discourse salience, the nature of this influence is likely to depend on the response type for a given task. This interaction with response type is considered in the following sections.

Task demands

With an experimental task, children's grammatical knowledge is inferred from explicit behavioral response, which maps onto a particular interpretation. As discussed above, this interpretation may be influenced by extragrammatical factors, including factors which affect parsing processes (e.g. interference and retrieval cost) or discourse salience. For an experimental context, an additional set of extragrammatical factors is the demands involved in generating a behavioral response. These *task demands* may interact with other extragrammatical factors, but may also vary depending on the task.

Tasks in previous studies may be broadly categorized into two types: first, tasks

which prompt participants to select a single interpretation from a set of options are *preference* tasks. In isolation, these tasks may identify interpretation preferences, but do not rule out disallowed interpretations. Next, tasks which include separate prompts for the adultlike and the non-adultlike interpretations are *judgment* tasks. By probing each interpretation, these tasks identify possible interpretations, but also impossible interpretations (Crain & Thornton 1998). Importantly, both types of task include task demands. However, the demands for these tasks may interact differently with other extragrammatical factors, due to the different response types.

Preference tasks. Preference tasks used with adjunct control include the Act Out Task (Goodluck 1981; Hsu, Cairns & Fiengo 1985; McDaniel, Cairns & Hsu 1991; Goodluck & Behne 1992; Cairns et al. 1994; Goodluck 1998; Goodluck 2001), Picture Selection Task (Janke & Bailey 2017; Janke 2018; Janke 2016), and Coloring Book task (Gerard et al. 2017; Gerard et al. 2018). In each of these, children's interpretation of PRO is inferred based on a selection between multiple choices, with each choice corresponding to a possible antecedent of PRO. Therefore, similar demands may be observed across tasks, based the procedures involved in making this selection. For example, task demands are imposed by interference from incorrect options, which must be avoided in order to select the correct option (Dillon et al. 2013; Jäger, Engelmann & Vasisht 2017). Meanwhile, other task demands will depend on the choices themselves, accounting for some of the variation between tasks.

In the Act Out Task, different possible referents (typically the main clause subject, object, or an external referent) are represented by toy characters, one of which is selected for acting out the adjunct clause for sentences like in (9), adapted from Cairns et al (1994):

(9) Ernie kisses Cookie Monster before PRO jumping over the fence

Either before or after this, two characters are also selected to act out the main clause (in (9), Ernie and Cookie Monster). Task demands are introduced by selecting each character, which requires avoiding any distraction from the other characters. These demands are increased if a non-target character is more salient. The choice of antecedent may also be influenced by the cost of acting out the main clause: if the adjunct clause is acted out first, then the main clause must be stored in memory while the adjunct clause is acted out; otherwise, the adjunct clause must be stored in memory while the main clause is acted out, which may affect the representation of the adjunct clause when it is retrieved. Importantly, the relevant contrast between an adultlike and a non-adultlike interpretation is between possible referents – for example, in (9), children’s behavior is adultlike or non-adultlike depending on whether they select Ernie (adultlike) or Cookie Monster (non-adultlike) to act out the adjunct clause.

Different contrasts are used in the Picture Selection Task and Coloring Book task, which involve different task demands (Pinto & Zuckerman 2018). In the Picture Selection Task, the relevant choice is between pictures which depict the events in the test sentence. For example, the correct picture for (9) would include the main clause event and Ernie jumping over the fence, while the incorrect picture would include the same main clause event but Cookie Monster jumping over the fence. Thus, in addition to the incorrect referent, the incorrect picture also depicts the same actions as in the correct picture, both of which must be ignored in order to select the correct picture (Adani 2011; Özge, Küntay & Snedeker 2019).

Finally, in the Coloring Book task, children are presented with a single black and white picture depicting both clauses, and prompted to select an item in the picture to color in, as in (10), adapted from Gerard et al (2018):

(10) Dora washed Diego before PRO eating the red apple.

In (10), children select between an apple held by Dora (adultlike) and an apple held by Diego (non-adultlike) to color in red, indicating their interpretation of PRO. Instead of the incorrect referent, as in the Act Out Task, children must inhibit the incorrect item (Diego's apple in (10)).

In general, the different contrasts across the three tasks are likely to impose different task demands. Meanwhile, unlike in the Act Out Task, neither the Coloring Book task nor the Picture Selection Task require an additional action for the main clause. Thus, in addition to the type of contrast which distinguishes between adultlike and non-adultlike interpretations, other differences (e.g. whether an action is required for the main clause) may introduce additional variation in the demands across tasks. As a result, comparisons cannot be made between the overall complexity of different tasks. Rather, by isolating and manipulating individual factors within tasks, the influence of these factors – e.g. parsing factors (Gerard et al. 2017) and discourse factors (Janke & Bailey 2017; Janke 2018) – can be identified for children's interpretations of adjunct control.

For grammars which generate more than one of the possible choices for selection, (e.g. both Ernie and Cookie Monster as possible antecedents in (9)), preference tasks require some interpretation strategy for making a selection. Moreover, this strategy may generate adultlike behavior, depending on the context. For example, an agent strategy for sentences with an active main clause predicts the same behavior as the adult grammar, as does a topic strategy (Goodluck & Behne 1992; Goodluck 1998; Goodluck 2001; Wexler 1992). Therefore, the influence of extragrammatical factors on children's responses with the adult grammar may be similar to the influence with a non-adult grammar, if an interpretation strategy is used which also generates adultlike behavior. In contrast, predictions diverge for adultlike and non-adultlike grammars in

judgment tasks, which are designed to identify *unavailable* interpretations, in addition to available interpretations.

Judgment tasks. Judgment tasks used with adjunct control include the reference judgement task (McDaniel, Cairns & Hsu 1991; Cairns et al. 1994) and the Truth Value Judgment Task (TVJT; Crain & Thornton 1998; Broihier & Wexler 1995; Adler 2006; Gerard et al. 2018). These tasks involve judgments about possible antecedents for PRO: in the judgment of reference task, these are explicit judgments based on metalinguistic reflection of possible referents, while the judgments in a TVJT are based on an experimental context, with different judgments corresponding to different interpretations of PRO. Like for the preference tasks, task demands may be observed in judgment tasks due to interference from incorrect options, and these demands will also vary depending on other aspects of the task. However, less variation has been observed between judgment tasks (29-43% of children with adult grammars) than between preference tasks (14-92% subject interpretations; Table 1).⁵ This suggests that there have been similar task demands across studies, or that overall task demands have involved similar complexity. For example, both tasks require a comparison between the

⁵ Higher proportions of subject interpretations are observed in the TVJT by Adler (2006); however, in this study the non-adultlike interpretation was always external, with sentences like in (iii), in contrast with other judgment studies:

(iii) John ate an apple before PRO running to the store.

With only one plausible antecedent for PRO in this sentence, any interference to this antecedent would be from an external referent, rather than an internal referent (as in (7)), which is likely to result in different task demands. Moreover, any preference for an internal referent will predict an adultlike pattern of responses, independent of the adult grammar (Goodluck, 1987; Cairns et al., 1994).

test sentence and the task context. Meanwhile, the relevant information for drawing this comparison may vary across tasks.

In the reference judgment task, children are prompted to judge whether sentences with adjunct control (and other anaphoric dependencies) are acceptable with a specific referent as the antecedent of PRO (McDaniel & Cairns 1990a; McDaniel & Cairns 1990b; McDaniel, Cairns & Hsu 1991; Cairns et al. 1994). By probing each referent individually, the reference judgment task identifies which interpretations are possible, but also for which interpretations are impossible.

To illustrate the intended meaning, the adjunct clause is acted out with toys by the experimenter. Therefore - similar to the Act Out Task - the reference judgment task involves no discourse context other than the toys in the acted-out demonstration. Also similar to the Act Out Task is the relevant contrast between an adultlike and non-adultlike response: a sentence is accepted if the referent presented for judgment is a possible antecedent, and rejected otherwise.

However, in contrast with the Act Out Task, judgment of a given referent will involve increased salience for that referent, relative to other referents in the discourse. To accept a grammatical referent, children must therefore avoid interference or other extragrammatical activation from any less salient ungrammatical referents in the discourse context. Meanwhile, rejecting an ungrammatical interpretation involves a more salient ungrammatical referent. As a result, children may be more likely to accept an ungrammatical referent than to reject a grammatical referent (see also Cairns et al. 2006). This contrast is supported by previous results with judgment studies, with a higher proportion of children accepting multiple referents in reference judgment studies than in TVJTs (Table 1).

In a TVJT, sentences with different possible interpretations are judged separately as true or false with respect to a context. Ideally, both true and false interpretations are equally salient in the context. Although children must identify a single interpretation for each sentence, it is assumed under the Principle of Charity (Crain & Thornton 1998) that children whose grammars make both interpretations available will access the true interpretation.

In the tasks discussed above, the relevant contrast between an adultlike and a non-adultlike response is generally determined by the type of task:

- in the Act Out and reference judgment tasks, the difference between an adultlike and non-adultlike response depends on the specific referent for the adjunct subject
- in the Picture Selection Task, this difference also depends on the adjunct subject, but other similarities between the correct and incorrect pictures must be disregarded to focus on the relevant difference
- in the Coloring Book task, children select one of two items, which correspond to an adultlike or non-adultlike interpretation

These different contrasts are likely to involve different task demands, as discussed above. Meanwhile, these contrasts are not used in previous TVJTs on adjunct control. Instead, the difference between an adultlike and non-adultlike response has depended on the order of events in a given context (Broihier & Wexler 1995; Adler 2006; Gerard et al. 2018). For example, given a context with the events in (11), the sentence in (12) can refer to two different event orders, depending on the interpretation of PRO:

- (11) a. Diego puts on a jacket
b. Dora tags Diego
c. Dora puts on a jacket

(12) Dora tagged Diego after PRO putting on a jacket.

In (12), the adultlike (subject control) interpretation is that Dora got a jacket. This makes the sentence in (12) false, since the order of events is stated in the reverse order of the context in (11). However, with a non-adultlike (object control) interpretation of (12), the order of events matches the order in (11), making (12) true. Importantly, the truth of the sentence in (12) depends on the order of events; the order of events is determined by the interpretation of PRO. Thus – assuming that both interpretations are equally salient – children’s interpretation of PRO may be inferred from their judgment of the event order in (12).

The complexity of previous TVJTs depends on the task demands of the individual processes involved in producing a truth value judgment, based on the order of events. Therefore, as with the preference tasks, the overall complexity cannot be ranked with respect to other tasks without quantifying the complexity of these processes. An important difference between the TVJT and other previous tasks, however, is the contrast between an adultlike and non-adultlike interpretation, which is more flexible in the TVJT. In the above example, this contrast depends on the order of events, but other contrasts are may also be used for investigating children’s interpretations of adjunct control. For example, the contrast from the Coloring Book task between two items is also compatible with a TVJT, with a context like the following (adapted from (10) above):

- (13) a. Dora washes Diego
b. Dora eats a blue apple
c. Diego eats an orange apple

(14) Dora washed Diego after PRO eating the orange apple.

In (14), the adultlike (subject control) interpretation is that Dora ate the orange apple. Like for the event order example in (12), this interpretation makes the sentence false in the corresponding context, since Dora's apple was blue. Meanwhile, the non-adultlike (object control) interpretation – that Diego ate the orange apple – is true in the context in (13).

Similar to the sentence in (12), these truth values are an indirect representation of PRO: in (14), the truth of the sentence depends on the color of the apple; whether this color matches depends on the interpretation of PRO. However, in contrast with (12), the truth values depend on the color of the apple, i.e. a property of an item in the adjunct clause, rather than the order of the events in both clauses.

Like the preference tasks, these TVJT examples have different contrasts between an adultlike response and a non-adultlike response. However, unlike the preference tasks, the response type in both TVJTs is the same, i.e. a true/false response as an indirect measure of children's interpretation of PRO. Therefore, a more direct comparison may be made between the task demands in these examples, depending on the contrast between an adultlike and non-adultlike response.

In previous TVJTs on adjunct control, only the event order contrast has been used. Accordingly, these TVJTs observed similar proportions of children with an adultlike pattern of behavior (Broihier & Wexler 1995; Gerard et al. 2018). Importantly, grammatical and extragrammatical accounts make different predictions for these proportions in a context with different task demands. These predictions are tested in the current study.

Predictions

For a TVJT with an event order contrast (an *event order TVJT*), a test sentence like (12) is compared to a context like (11). After hearing the story context and representing this order of events, children must parse the test sentence in (12) and retrieve the antecedent of PRO, with potential interference from other referents in the discourse context as discussed above. Next, with the relevant event in the adjunct clause determined by the antecedent of PRO, both the main clause and adjunct clause events in (12) must be maintained in memory, in order to compare the ordering of these events with the order in the story context in (11). A true/false response is given based on whether these orders match.

This contrast involves much more information than a TVJT that depends on the color (a property) of an object in the adjunct clause (a *coloring TVJT*), as in (13) and (14). Just as in the event order TVJT, children must parse the test sentence and retrieve the antecedent of PRO, introducing similar parsing demands. However, with the relevant item in the adjunct clause determined by the antecedent of PRO, only this item is needed – in (14), Dora or Diego’s apple – to compare the color of the item in the sentence with the relevant item in the context. A true/false response is given based on whether these colors match.

To produce a true/false response, the event order TVJT requires information from both clauses and the relation between clauses, while the coloring TVJT requires only the item in the adjunct clause and a basic property of this item (specifically, its color). As a result, higher task demands are expected for the event order TVJT than for the coloring TJVT. These differences may be realized in the process of comparing the test sentence to the context, or even as the test sentence is encoded and maintained in memory, e.g. if the need to encode more detailed information about the events in both

clauses restricts the resources available for antecedent retrieval. Based on these differences, the event order TVJT is defined for the current study as *high demand* relative to the coloring TVJT, which is therefore defined as *low demand*. For this difference in task demands, contrasting predictions are made for grammatical and extragrammatical accounts of children's non-adultlike behavior.

The sentences in (12) and (14) are false under a subject control interpretation and true under an object control interpretation. Therefore, children with the adult grammar (and adults) should reject these sentences in the contexts in (11) and (13), respectively. Meanwhile, a free reference grammar generates both interpretations; therefore, under the Principle of Charity (Crain & Thornton 1998), children with a free reference grammar should select the (true) object control (true) interpretation over the (false) subject control interpretation.

These patterns of behavior are modulated by the task demands, however. Importantly, children's behavior in a **low** demand context will more closely reflect the interpretations generated by their respective grammar, while a high demand context is more likely to introduce extragrammatical factors which influence children's behavior. Specifically, for sentences which are false under a subject control interpretation but true under an object control interpretation:

- Children with the adult grammar should reject these sentences – therefore, a **lower** acceptance rate is expected in a low demand context.
- Children with a free reference grammar should accept sentences (due to the Principle of Charity) – therefore, a **higher** acceptance rate is expected in a low demand context.

These predicted patterns may be compared with children's behavior for sentences which are true under a subject control and false under an object control interpretation, which make the same prediction for both grammars:

- Children with the adult grammar or a free reference grammar should accept these sentences – therefore, a **higher** acceptance rate is expected in a low demand context.

In both sets of predictions, the low demand is defined as low relative to the high demand task; thus, children's behavior on the low demand task must be interpreted relative to the high demand task. That is, the low demand task is not a “no demand” task. Moreover, for a low demand task, the predicted acceptance rates are not at 0% or 100% for the non-adult grammar or the adult grammar. Recall that all tasks have demands, and - especially for complex sentences - children's average accuracy rarely if ever reaches 100% (Hamburger & Crain 1984; Conroy et al. 2009; Crain & Thornton 1998; Crain et al. 1996; Drozd 2001). That is, less than 100% accuracy should not be taken as evidence against an adult grammar, because children's behavior with an adult grammar is not generally expected to reach 100% accuracy. Rather, children's accuracy is expected to improve in a low demand task relative to a high demand task: sentences that are not permitted in the grammar are more likely to be rejected, and sentences that are permitted that are more likely to be accepted.

The current study tests these predictions in the following section. Consistent with the adult grammar, in the low demand task children accepted more subject sentences and fewer object sentences than in the high demand task. This suggests that extragrammatical factors play a considerable role in children's acquisition of adjunct control. Implications are discussed in the final section.

Experiment

In this section the above predictions for different task demands are tested in two TVJTs. Children's behavior in a task with high demands is compared with a task with low demands. Specifically, the high demand task required children to remember the order of events in a story as in the event order TVJT described above, while the low demand task involved the color of an object, as in the coloring TVJT. The implications of these differences are discussed further below.

Participants

Participants for the high demand TVJT with an event order contrast were 32 children (13 males) ages 4;0-5;6 ($M=4;9$). An additional 22 children were excluded from the sample for answering too many filler sentences incorrectly (18) or failure to complete the training portion (4). For the low demand TVJT with a color contrast, another 32 participants were recruited from the same general area and with similar demographics (15 males), ages 4;0-5;2 ($M = 4;6$). An additional 22 children were excluded from this sample for answering too many filler sentences incorrectly (21) or inattention (1). Both groups were recruited through the University of Maryland Infant and Child Studies Database or participated at their local preschools.

To confirm the expected responses outlined above, adult controls were also tested on the event order TVJT ($n=32$) and the coloring TVJT ($n=32$). Two additional adults were excluded (1 from each sample) for answering too many filler sentences incorrectly. The adults were monolingual native speakers of American English recruited from the online experiment platform Prolific, and were compensated at a rate of \$9.50 per hour.

This study was approved by the Institutional Review Board at the University of Maryland, College Park (for children) and the Ulster University Risk and Ethics


Committee (for adults), with written informed consent collected from all adult participants and parents of child participants.

Design and procedure: high demand

The high demand task was the event order TVJT described above, based on previous studies on the acquisition of adjunct control which have used event order contrast as the relevant contrast between the true and false interpretations (Broihier & Wexler 1995; Adler 2006; Gerard et al. 2018).

Short stories set up contexts which made both a subject control and an object control interpretation available. Test sentences either made the subject interpretation true and the object interpretation false (*subject control* condition) or the subject interpretation false and the object interpretation true (*object control* condition), manipulated within-subjects. An example story is presented in Table 2, with a verbatim script and pictures that were presented with each line (All stories are presented in Appendix A). The script was performed by the experimenter, with different voices for each character (narration, Dora, and Diego). Characters were animated to wiggle while they spoke, and actions were also animated (e.g. running, getting a jacket).

Table 2: a context and test sentences for the event order TVJT.

Animation	Script
	<p>Narrator: Diego and Dora are going outside to play in the snow. It’s so cold outside! Diego is wearing a hat, but that’s not enough.</p>



Diego: I'm gonna go get a jacket so I won't be so cold anymore. Dora do you want a jacket too?

Dora: No thanks, I don't want a jacket now.

Diego: But it's so cold outside! How will you stay warm?

Dora: I won't be cold if we play tag. Then we can run around and we won't get cold.

Diego: Ok, but I'm still gonna get my jacket, it's too cold out here.



Narrator: So Diego gets his jacket, and then they start playing tag.



Narrator: Dora is it, and she chases Diego around in the snow.



Diego: I bet Dora won't find me if I hide behind this snowman!

Dora: I saw Diego hide behind that snowman! I'm gonna tag him!



Dora: Tag, you're it!

Narrator: And Dora tags Diego so hard that they both fall down and get covered with snow.



Dora: Hmm, it is pretty cold out here. You're covered in snow, Diego, aren't you cold?

Diego: Nope, I've got a jacket already so I'm not cold at all!

Dora: Oh wow, I'm so cold! What can I do so I won't be cold anymore either?

Diego: Why don't you get a jacket too? Then you won't be so cold anymore.

Dora: Oh, what a great idea! I'll go get my jacket too.



Dora: Look Diego, I got a jacket! So now we can play tag again. You're it!



- (15) *Preamble:*
Dora and Diego were both
playing tag outside, and...



- (16) a. *Subject control:*
Dora tagged Diego before getting a
jacket
b. *Object control:*
Dora tagged Diego after getting
a jacket

For the test sentences in (16), the adultlike interpretation is that Dora got a jacket, which makes the *subject control* sentence true and the *object control* sentence false. The non-adultlike interpretation is that Diego got a jacket, which has the opposite truth values *because* this describes the reverse order of events in the story. Thus, in order to correctly judge the sentences, children needed to be able to retrieve an antecedent of PRO *and* judge whether the events described in the test sentence were stated in the correct order.

In addition, to balance the salience of each referent before the test sentence, a preamble (as in (15)) was delivered just before the test sentence. The preamble was a short sentence about the story, and was delivered by the puppet (Crain & Thornton 1998). In the preamble, neither Dora nor Diego is a clear topic (that is, the preamble did not provide a pragmatic lead for one of the referents), and the order of mention in the conjoined clause was also counterbalanced across lists items and lists.

Children were also prompted to briefly summarize the main events in the story right before the test sentence was delivered, and to give justifications for their true/false answers. To help with these, images of the main events from the story remained visible from the end of the story until the beginning of the next trial.

A training session ensured that children could correctly judge simple sentences with *before* and *after* (Appendix A), and children who did not do so despite corrective feedback did not proceed to the testing portion. Filler trials⁶ were also interspersed with the test trials to confirm that children could judge minimally different sentences whose truth value did depend on the order of events, but which did not involve antecedent retrieval:

(17) Dora hugged Diego before/after the plane landed.

(where Diego has arrived on a plane to meet Dora)

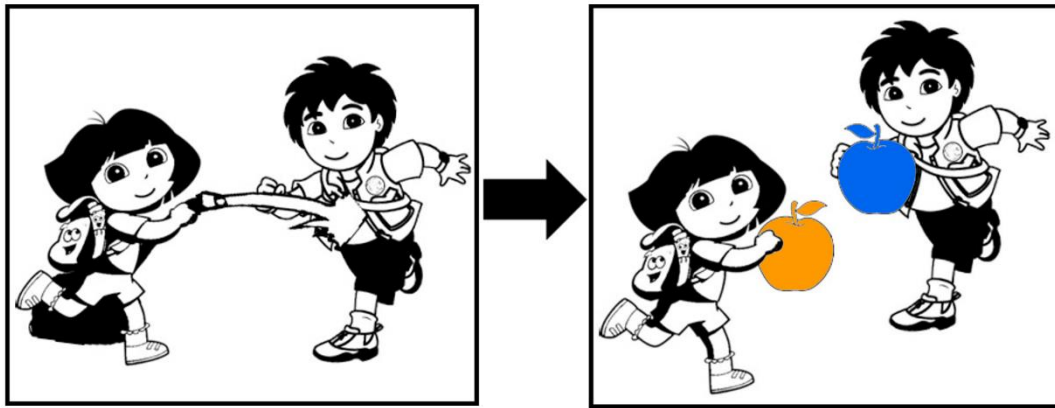
Training and filler sentences involved temporal relations, but contained no features of syntactic control. In total, children saw six training items with and without visuals, four test items as in (16), and three filler items as in (17). Children were included in the analysis if they answered zero or one filler question incorrectly, but were excluded if they answered more than one filler question incorrectly. Stories were presented with PowerPoint on a touchscreen PC, with sessions lasting from 20-25 minutes for children and 10 minutes for adults.

Design and procedure: low demand

The low demand task was the coloring TVJT described above, which was adapted from the Coloring Book task (Gerard et al. 2018; Pinto & Zuckerman 2018). Instead of short stories as in the event order TJVT, children saw static pictures as in Table 3, with characters from the test sentences in (19):

⁶ Technically these are control trials and not fillers, as they are used to define exclusion criteria. However, they are referred to here as fillers to avoid confusion with the test trials with adjunct control.

Table 3: a context and test sentences for the coloring TVJT.



In this picture, we have we have Dora washing Diego, and then there's Dora eating an apple, and there's Diego eating an apple too.

(18) *Preamble:*

Ok [puppet], can you tell us how [friend] colored this picture of Diego and Dora?

(19) a. *Subject control:*

Dora washed Diego before eating the orange apple.

b. *Object control:*

Dora washed Diego before eating the blue apple.

One picture showed the characters from the test sentence performing the action in the main clause (*Dora washed Diego*), and the other picture showed both characters performing the action from the adjunct clause (*eating the apple*). To allow for a contrast between referents, two of the items were also filled in with different colors. For example, the apples in Table 3 were filled in so that Dora's apple was orange and Diego's apple was blue (unconventional colors were used to avoid judgments based on real-world biases, with different colors for different items). This color contrast distinguished the two conditions, in (19).

For the sentences in (19), the adultlike (subject control) interpretation is that Dora ate the apple. This makes the sentence true in (19a), because Dora's apple is orange. The truth value depends on the interpretation of PRO however: the object

control interpretation is that Diego ate the apple, which makes (19a) false because Diego's apple is blue (with the reverse truth values in (19b)).

Similar to the high demand TVJT, a preamble (as in (18)) was delivered before the test sentence to balance the salience of each referent (Crain & Thornton 1998). As with the high demand TVJT, both referents were mentioned in a conjunct, meaning that neither Dora nor Diego was a clear topic for the test sentences. The order of mention in the conjoined clause was also counterbalanced across lists items and lists.

In order to judge the test sentences, children needed to compare the color in the test sentence to the color in the picture. This meant retrieving the antecedent of PRO from the main clause to identify which apple's color should be mentioned in the test sentence, and judging whether this color matched the relevant apple in the picture. Additionally, to confirm that both the subject interpretation and the object interpretation were available (and to prevent children from developing a strategy of focusing on just one part of the sentence), we included filler sentences with a finite adjunct and an overt pronoun subject:⁷

⁷ Different filler items were used in the coloring TVJT and event order TVJT to reflect the relevant contrast between a true and false response, which children must be sensitive to in order to give a meaningful truth value response. In the event order TVJT, this meant that filler items were designed to confirm that children could accept or reject a sentence depending on the order of events. Meanwhile, the filler items in the coloring TVJT had a different structure, to confirm that children could judge a sentence based on the color of the item in the adjunct clause, and that this judgment was available for both a subject and an object interpretation of the adjunct subject.

(20) a. Dora washed Diego before she ate the orange apple

(subject antecedent, true)

b. Dora washed Diego before she ate the blue apple

(subject antecedent, false)

(21) a. Dora washed Diego before he ate the blue apple

(object antecedent, true)

b. Dora washed Diego before he ate the orange apple

(object antecedent, false)

Filler items were interspersed with test items, and were balanced for the number and order of subject and object pronoun antecedents. Truth values were set dynamically, based on the participant's previous answer, to avoid answer biases; e.g. if a participant had just previously answered *true*, then for the following filler item the false version of the sentence would be delivered (and vice versa). As in the other TVJT, children were prompted to give justifications for their true/false answers.

A training session ensured that children could judge sentences with colors correctly. In total, children saw five training items, four test items as in (19), four filler items with a pronoun subject antecedent as in (20), and four filler items with a pronoun object antecedent as in (21). Children who responded incorrectly to more than one filler item with a subject pronoun antecedent or to more than one item with an object pronoun antecedent were excluded from the analysis. That is, children could answer up to two filler items incorrectly (one of (20) and one of (21)) and still be included in the analysis. The pictures were presented with PowerPoint on a touchscreen PC, with sessions lasting around 15 minutes for children and 5 minutes for adults.

Justifications

In addition to the true/false responses, children gave justifications for their answers, usually with little prompting (Table 4). Children's explanations were transcribed and coded for whether their justification made it clear who they had interpreted as the antecedent of PRO, and if not, whether their justification was relevant to the task context. Clear justifications were coded as 'clear,' and for which character had been interpreted as PRO (Dora or Diego). Unclear justifications were either coded as 'unclear,' but still relevant to the task context, or 'irrelevant,' if the child forgot their answer or was unable to give a justification (Syrett & Lidz 2011).

Table 4. Breakdown of justifications by clear, unclear, and irrelevant (raw numbers).

		Code		
Task		Clear	Unclear	Irrelevant
Task	high demand (event order)	110 (54 Diego, 56 Dora)	11	7
	low demand (color)	104 (48 Diego, 56 Dora)	18	5

Children's justifications were not different from each other across the two tasks, and the responses patterned similarly whether or not the unclear responses are included.

Results and discussion

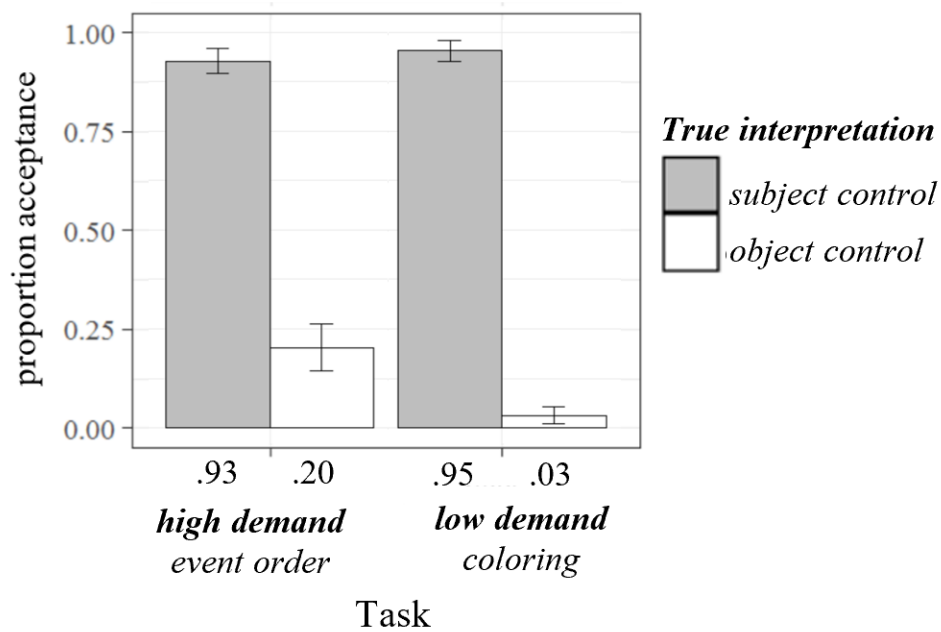
This section first reviews the results from adults, followed by the results from children.

Adults

Results for both tasks with adults are presented in Figure 1. To analyze these results, we used R (R Core Team 2015) and lme4 (Bates et al. 2015) to perform a mixed-effects logistic regression, with Task and True interpretation (subject control/object control) as contrast-coded fixed effects, and subjects and items as random effects. The maximal

model (random intercepts only, since a model with random slopes did not converge (Barr 2013; Eager & Roy 2017)) revealed a main effect of True interpretation ($\beta = 3.8$, $Z = 5.99$, $p < .001$), a main effect of Task ($\beta = -2.27$, $Z = -2.9$, $p = 0.004$), and a significant interaction between True interpretation and Task ($\beta = 2.78$, $Z = 2.56$, $p = .01$).

Figure 1: Effects of true interpretation and task demand - depending on order of events (high demand) or a contrast in color (low demand) - on sentence acceptance in adults.



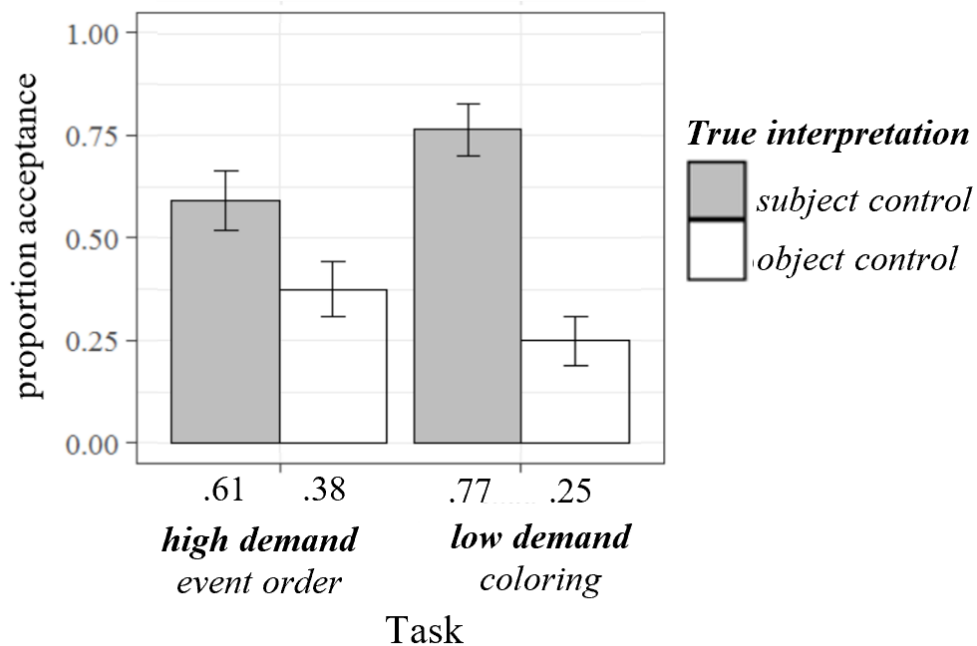
The main effect of True interpretation in Figure 1 reflects that adults accepted more *subject-control* sentences than *object control* sentences overall; this was observed for both low and high demand tasks. Meanwhile, the main effect of Task is due to the higher overall acceptance rate in the high demand task. However, this effect is driven by the acceptance rate in the *object control* condition of the high demand task, which is reflected by the interaction between True interpretation and Task. Importantly, a greater effect of True interpretation is observed for the coloring TJVT than for the event order

TVJT, suggesting that even adults – whose grammars only generate a subject control interpretation in a neutral discourse context – were influenced by the task demands in the event order TVJT. This difference is also observed in the distribution of responses, with a strict subject control pattern observed in 18 of 25 adults in the event order TVJT, but from 22 of 25 adults in the coloring TVJT. Therefore, the same difference between tasks is expected from children with the adult grammar (although with lower proportions of subject control overall, due to children’s greater sensitivity to the task demands, as discussed above).

Children

Results for both tasks with children are presented in Figure 2. As for the adults, we used R (R Core Team 2015) and lme4 (Bates et al. 2015) to perform a mixed-effects logistic regression, with Task and True interpretation (subject control/object control) as contrast-coded fixed effects, and subjects and items as random effects. The maximal model (random intercepts only, since a model with random slopes again did not converge (Barr 2013; Eager & Roy 2017)) revealed a main effect of True interpretation ($\beta = 1.24$, $Z = 2.93$, $p = .003$) and a significant interaction between True interpretation and Task ($\beta = 1.58$, $Z = 2.49$, $p = .01$), with no effect of Task ($\beta = -0.72$, $Z = -1.41$, $p = 0.16$).

Figure 2: Effects of true interpretation and task demand - depending on order of events (high demand) or a contrast in color (low demand) - on sentence acceptance in children.



The main effect of True interpretation reflects that children accepted more *subject control* responses than *object control* responses overall; this was observed for both low and high demand tasks. Meanwhile, the interaction is due to the greater effect of True interpretation for the low demand coloring TVJT, just like for adults: in the coloring TVJT, children accepted more subject-true sentences *and* accepted fewer object-true sentences. That is, in the low demand task, children's grammars more closely resemble the adult grammar. This difference is also observed in the distribution of responses (Table 5):

Table 5: Distribution of children's responses across tasks.
(*true* = *subject control* condition, *false* = *object control* condition)

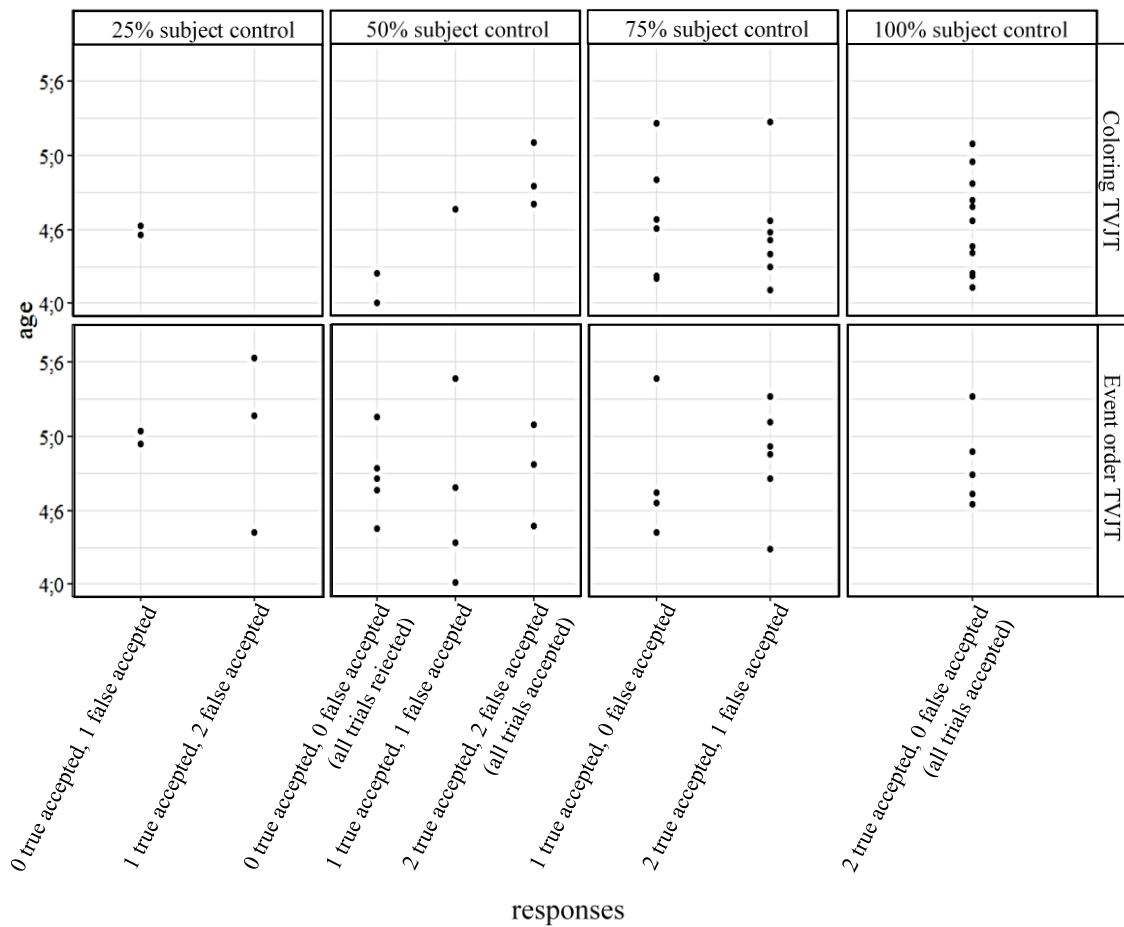
Response pattern	Event order TVJT	Coloring TVJT
2 true accepted, 0 false accepted (adult-like)	5	11
2 true accepted, 1 false accepted	6	7
1 true accepted, 0 false accepted	4	6
2 true accepted, 2 false accepted (all trials accepted)	3	3
1 true accepted, 1 false accepted	4	1
0 true accepted, 0 false accepted (all trials rejected)	5	2
1 true accepted, 2 false accepted	3	0
0 true accepted, 1 false accepted	2	2

In Table 5, strict subject control is realized as accepting both *subject control* (*true*) responses and rejecting both *object control* (*false*) responses. To compare children's responses with the previous studies in Table, a pattern of 75% subject control responses also includes a response pattern with both *subject control* responses accepted and one *object control* response accepted, as well as both *object control* responses rejected and one *subject control* response accepted. These patterns are highlighted in gray. Of the 32 children in each study, 15 children (47%) exhibited this pattern in the event order TVJT, while 24 children (75%) exhibited this pattern in the coloring TVJT.

Finally, children's responses in both tasks were not predicted by age, with all response patterns in Table 5 observed across all ages (Figure 3). This is consistent with previous studies: with the exception of longitudinal studies (McDaniel, Cairns & Hsu 1991; Cairns et al. 1994), age effects are not observed across this age range (Hsu, Cairns & Fiengo 1985; Goodluck & Behne 1992; Goodluck 1998; Gerard et al. 2017; Gerard et al. 2018).⁸

⁸ Lower accuracy is generally observed for adjunct control in 3-year-olds, with higher accuracy in 7 and 8-year-olds; however, age effects are consistently absent in 4-6 year-olds in previous studies on adjunct control, in cross-sectional designs (Hsu et al., 1985; Lust et al., 1986; Gerard, 2016).

Figure 3: Response patterns in both tasks by children's age.



Discussion

The above results are consistent with the predictions of the adult grammar, but not with the predictions of a free reference grammar. This suggests that children's grammars are adultlike, but their interpretations may be affected by extragrammatical factors, particularly in a high demand context. Therefore, these factors are also likely to have influenced children's behavior in previous studies, especially for tasks that were similar to the high demand TVJT. For example, previous TVJTs also used the order of events as the primary difference between subject and object interpretations (Broihier & Wexler 1995; Adler 2006). One consideration is that previous studies have tested a wider range of complementizers than the current task, which only used *before* and *after*. In sentences

with *before* and *after*, separate events must be held in memory, in addition to their relative ordering. Previous studies have included other complementizers like *while*, which also involves multiple events but without a sequential ordering relation. In both an act out task (Cairns et al. 1994) and a TVJT (Adler 2006), children's responses for *while* mirrored their responses for *before* and *after*. This presents two possibilities: first, *while* does not involve an ordering relation (in a specific sequence) but does still involve a temporal relation (i.e. a relation between the timing of events); thus, the observed difficulty may be due to the temporal nature of the relation between clauses which is present for all three complementizers (*before*, *after*, and *while*). Another property of all three complementizers is the relevance of multiple events; children's difficulty may therefore stem from having to hold multiple events in memory rather than computing the relation between them. Finally, both factors (relative timing and multiple events) may contribute to children's difficulties.

Meanwhile, children's improved accuracy in the low demand TVJT raises questions about the source of children's improvement, as well as the persisting task demands. Based on the predictions discussed above regarding the amount of information involved in the different types of contrasts, it is expected that the cost of comparing event orders should be greater than comparing colors. It is less intuitive exactly how this difference should influence antecedent retrieval in sentences with adjunct control or if it should affect truth value judgments. Also, despite the increase in accuracy which was consistent with the predictions of the adult grammar, children's responses were still not at ceiling for the low demand task, with some object interpretations remaining in both conditions (subject-true and object-true). As described above, these residual object responses are still consistent with an adult grammar, since the low demand task was exactly that - while the demands were lower, they were not

absent. Thus, children's responses were a more accurate reflection of their grammar than for the high demand task, but they were still influenced by the context of the low demand task. In general, understanding how these demands can influence children's interpretations will help to clarify the role of the linguistic input for acquiring adjunct control, particularly if similar demands are present in the input. These implications for children's acquisition are discussed further in the following sections.

General discussion

In this paper we asked about grammatical and extragrammatical factors in children's non-adultlike interpretations. Children's behavior for sentences with adjunct control was compared on two TVJT tasks: a high demand task, and a low demand task. In the high demand task, the relevant contrast between the adultlike and non-adultlike interpretation involved a more complex computation involving the order of events, while the same contrast in the low demand task depended on an item's color. Children's behavior was more adultlike when the judgement involved a contrast in color compared to event order, suggesting that the complexity of the contrast can affect the deployment of linguistic knowledge. Moreover, in previous tasks with comparable complexity, children's behavior was likely to have been affected in similar ways. This is discussed further in the following sections, which relate the present results to children's behavior in previous studies, participants in the current study compared to previous studies, and further implications for acquisition.

Results compared with previous studies

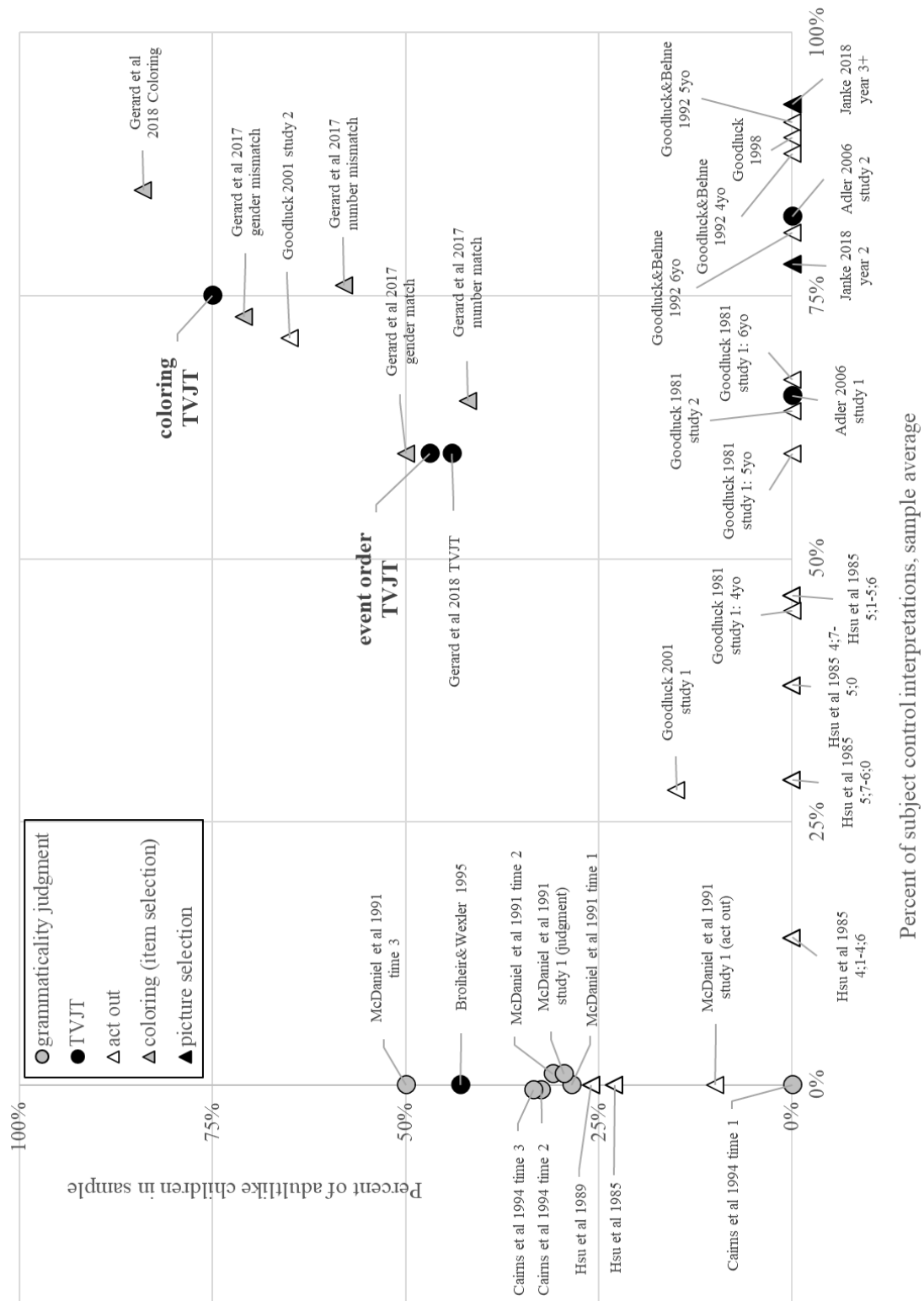
In previous studies on the acquisition of adjunct control, different patterns of non-adultlike responses have been observed across different tasks. These tasks differed in a number of ways, as discussed above: in addition to the type of measure (act out,

judgment), different steps were involved in linking the sentence to the task context to generate a response. However, the different patterns have generally been attributed to differences in the proportion of non-adult grammars in the subject population, rather than the task.

To compare the results from the current study to previous results, both tasks are plotted in Figure 4 with previous studies, based on the measures reported in Table 1 of the proportion subject interpretations in the sample, and the proportion of adultlike children (with children who gave at least 75% subject interpretations categorized as adultlike).

For the event order TVJT, children's pattern of responses is strikingly similar to the patterns observed in previous TVJTs – in particular, the same proportion of children with an adultlike pattern of behavior was observed as in other TVJTs with a similar design (Broihier & Wexler 1995; Gerard et al. 2018), and the same proportion of subject responses was observed in the sample as in Gerard et al (2018), which used the same materials. Therefore, although two separate samples were tested on the event order TVJT and coloring TVJT, children's similar behavior in the event order TVJT to previous TVJTs suggests that these samples are comparable. In contrast to the more adultlike behavior on the coloring TVJT, this consistency between the event order TVJT and previous TVJTs supports the interpretation of the current results as due to the task manipulation, without requiring a within-subjects manipulation.

Figure 4: Response patterns in previous studies, by proportion of subject control interpretations in the sample, and proportion of adultlike children in the sample.



Meanwhile, the results of the current study are consistent with other within-task manipulations, which have also used different samples in different conditions. For example, in Gerard et al's (2017) manipulation of similarity-based interference, children exhibited more adultlike behavior in conditions with less interference. Similarly, Goodluck (2001) observed more adultlike behavior in a manipulation with two different sentence structures (Goodluck 2001: Study 2), compared to a design with three different structures (Goodluck 2001: Study 1). These effects suggest that children's behavior is influenced by different extragrammatical factors, in addition to the different task demands tested in the current study. That is, in addition to the relevant contrast between the adultlike and non-adultlike response, children's interpretations of adjunct control may also be influenced by other extragrammatical factors, as discussed above. Moreover, these factors may also account for children's residual non-adultlike responses.

Residual object control

Although children's responses were more adultlike in the coloring TJVT, the pattern of responses still included non-adultlike behavior at a rate of 24% across conditions. One possibility is that these non-adultlike responses were due to a non-adult grammar; however, the responses in the coloring TVJT must be interpreted relative to the responses in the event order TVJT. Although the manipulation in the current study was between-subjects, the similar pattern of behavior in the event order TVJT and previous TVJTs suggests that the distribution of children's grammars does not vary significantly across samples. Therefore, the distribution of grammars alone cannot account for the difference between tasks in the current study.

As a result, any grammatical account of children's responses in the current study must explain the more adultlike pattern of behavior observed in the coloring TVJT. This difference between tasks is not predicted by a free reference grammar: as discussed above, children's responses in the coloring task should more closely resemble their grammatical knowledge, which for a free reference grammar translates to a *higher* proportion of responses accepted in the *object control* condition for the coloring TVJT. As this prediction was not borne out, children's behavior is not consistent with a free reference grammar. Moreover, since the current study tested an object control interpretation, but not an external interpretation, the same conclusion applies to an optional subject-object grammar (which also predicts an increase in object control responses in a low demand task). If the difference between the event order and coloring TVJTs is not predicted by a non-adult grammar, then this difference must be due to extragrammatical factors.

At the same time, several participants in the current study were excluded in both experiments. Were their grammars non-adultlike? Or, put another way, would they have been responsible for the non-adultlike responses as participants in previous studies?

Excluded participants

Both of the tasks in the current study were TVJTs, meaning that children's responses revealed their interpretations of PRO indirectly based on a *true* or *false* response. Importantly, their follow-up justifications generally aligned with their initial responses, suggesting that the true/false answers were a relatively valid measure of children's interpretations. Complications arise when considering the true/false responses in the contexts of other tasks, however, due to the different strategies that may be observed in high demand contexts. For example, many of the children who were not included in the

analysis of the current study exhibited an answer bias; i.e. persisted in giving the same true/false answer throughout, despite giving both *true* and *false* answers throughout the initial training session (Table 6). Due to the dynamic nature of the filler trials (the correct answer was always the opposite of the child’s response to the previous trial), these children answered the filler trials incorrectly.

Table 6. Number of subjects, by bias. Participants were not included if they answered 2 or more filler sentences incorrectly.

Task demand	Not biased, included	Not biased, not included	<i>True</i> bias	<i>False</i> bias
High	32	8	9	5
Low	32	11	7	4

The tendency for a *true* bias is not generally expected from children at this age, although previous research on children’s biases has generally focused on less complex sentences (Fritzley, Lindsay & Lee 2013; Fritzley & Lee 2003). Still, most of these children with answer biases gave clear justifications, and from these it was clear that most of the children were not ignoring the test sentences. Thus, many of these children’s answers to the test trials were entirely interpretable; however, based on their answers to the filler trials (which were incorrect), it could not be determined whether their interpretation of the test sentence was affected by their answer bias.

In general, the responses of the children with answer biases are not an accurate reflection of their grammars. Moreover, it is not entirely clear whether or how their responses would be influenced in a different task with a different response type. For example, if the answer bias reflects a general tendency to persevere on a particular response, then in an act out task children might choose the same antecedent for every

trial. However, previous act out tasks have generally used a range of characters, which would have prevented a consistent antecedent bias. Other task-specific biases are not discussed for previous act out tasks. Meanwhile, a similar *true/false* bias might be expected for a grammaticality judgment task, although these biases are not mentioned in previous judgment studies (McDaniel, Cairns & Hsu 1991; Cairns et al. 1994).

For the children with answer biases, there were minimal differences across the two tasks (high demand and low demand), with similar justifications given in both groups. In contrast, the children who did not pass the filler trials and did not have answer biases were different in the different tasks. In the high demand task, most of the children who did not pass the filler trials also did not pass the training session. That is, they were unable to judge more basic sentences with *before* and *after*. These children may not have learned the meanings of *before* and *after*, or they may have been overwhelmed by the demands of judging sentences with *before* and *after*. Either way, their responses in the high demand TVJT would not have reflected their grammatical knowledge of adjunct control. Additionally, their responses in and act out or judgment task may have been similarly affected, depending on the source of their difficulties (lexical or extragrammatical).

In contrast, none of the children in the low demand task failed the training portion, which also used sentences with *before* and *after*. Rather, these children either exhibited a subject bias overall (in particular, 4 children who answered the object pronoun fillers incorrectly), or misunderstood the task (e.g. one child had a “Dora” bias, and another judged sentences based on their own color preferences). Importantly, the fillers were important for identifying any children who were not responding based on their grammatical knowledge, either due to a response bias or otherwise. One possible exception is the children with a subject bias, who may either have the adult grammar for

adjunct control and a bias for subject pronouns, or a general subject bias. This raises the question of biases, however, which are potentially problematic for the results of the current study.

Subject response biases

The current study compares the predictions of an adult grammar (strict subject control) with non-adult grammars which also allow object control. The predictions for the adult grammar are that in a less demanding TVJT, children will accept more subject responses and fewer object responses. However, the same predictions are made for any non-adult grammar which allows multiple interpretations but which has a non-syntactic bias that usually picks out the subject (assuming that the bias overrules the Principle of Charity). For example, if children have a bias to interpret the adjunct subject as the discourse topic, then this will usually pick out the main clause subject as the antecedent of PRO. Similarly, if children have an agent bias, then this will pick out the main clause subject whenever the main clause is active, but not when the main clause is passive (Goodluck & Behne 1992; Goodluck 1998). Finally, if children have an animacy bias, then this will pick out the main clause subject as the antecedent of PRO whenever the main clause subject is animate, but no other internal NP is animate. In the current study with two animate NPs in the main clause, this last strategy would not lead to the observed results, although adults exhibit an animacy bias in addition to the syntactic control dependency, meaning that children do acquire this bias at some point (Parker, Lago & Phillips 2015). For the remaining biases (topic and agent), there are issues from an empirical perspective, as well as for learnability.

In typical conversation, the subject and topic align so that the topic will identify the grammatical antecedent of adjunct PRO. However, the current study (and previous studies in general) was balanced to make both referents equally salient, without one

being a clear topic over the other. This aim was achieved by making both characters equally relevant in the story, along with the preamble to balance the salience of each referent directly before the test sentence. Thus, if children have a topic bias, it cannot explain the results in the current study.⁹

Next, the results from the current study cannot rule out an agent bias. However, the results of studies which have tested an agent bias do not observe one: although Goodluck has observed fewer subject responses with a passive main clause than with an active main clause (i.e. more non-adultlike agent responses), children did not show a bias for agent responses over subject responses in an act out task (Goodluck & Behne 1992; Goodluck 1998). Moreover, Gerard (2016) observed significantly more subject

⁹ A reviewer suggests that children's responses may have been influenced specifically by a topic interpretation of the preamble used for the event order TVJT, repeated below as (iv):

(iv) Dora and Diego were both playing tag outside, and... [test sentence]

In particular, because children's (and adults') responses may be influenced by a pragmatic lead (Janke & Bailey, 2017; Janke, 2018) it is suggested that children's responses in the event order TVJT may have been influenced by topic information in the conjoined clause.

However, previous studies which have manipulated the topic have included just one of the possible referents in a pragmatic lead before the test sentence, making it unlikely that children's responses would be influenced by just one of the referents in the conjoined clause with both of them. Additionally, children's responses were not influenced by a pragmatic lead in previous studies at this age (4-6 years). Instead, influence due to a pragmatic lead is not likely to be observed until at least age 8 or older (Janke, 2018). Finally, since children and adults showed the same pattern of behavior in the current study (i.e. a higher proportion of subject control in the coloring TVJT), this pattern is likely to have the same source - that is, if the topic is not involved in determining children's interpretations, then this is unlikely to be the cause of the difference between tasks for adults - especially with both referents in the preamble instead of rather than just one (as in a pragmatic lead).

responses than agent responses in sentences with a passive main clause, using a TVJT. Assuming that children use the same antecedent retrieval procedures with active and passive main clauses, these results suggest that children do not use an agent strategy for identifying the antecedent of PRO.

More generally, biases which pick out a subject interpretation are problematic from a learnability perspective: if children need evidence from the linguistic input to transition from a non-adultlike grammar to the adult grammar, then they need a signal that their current grammar is not the correct one (generally by being unable to parse the input with their current grammar). However, biases which pick out a subject interpretation will generally pick out the same set of interpretations as the adult grammar. Exceptions that tease apart a subject bias and the adult grammar - e.g. where the subject is not the topic, or with a passive main clause - are rare, and even if they occur, they would not produce the kind of error signal that would be needed for a transition to the adult grammar. This is because the bias would still identify an antecedent, even an ungrammatical one. That is, if the input contains an instance of adjunct control with a passive main clause, then a child with an agent bias would get a non-adultlike interpretation. Based on patterns of behavior in previous studies, children are unlikely to realize if they get an ungrammatical interpretation, so this potential evidence would not be used for transitioning to the adult grammar. Thus, children with a subject bias would not be able to use the most relevant form of evidence to transition to the adult grammar. This suggests that children do not have a subject bias, and rather that their subject responses are due to the adult grammar. If so, an account is needed for how children acquire the adult grammar.

Acquiring the adult grammar

To acquire the adult grammar of adjunct control, children need the correct attachment height for the non-finite adjunct and the correct relation between the PRO and its antecedent - the closest c-commanding NP. Without either of these, non-adultlike interpretations are predicted. Correspondingly, to account for children's non-adultlike interpretations, non-adultlike grammars have been proposed which lack each of these components (Goodluck 1981; Hsu, Cairns & Fiengo 1985; McDaniel, Cairns & Hsu 1991; Cairns et al. 1994; Wexler 1992; Broihier & Wexler 1995; Goodluck 2001; Adler 2006). These accounts vary in their predictions for how children transition to the adult grammar, depending on what kind of evidence is needed for attachment height or the correct antecedent. Additionally, Wexler's (1992) Nominalization account involves the maturation of two separate grammatical components, independent of the input.

Importantly, if children access non-adultlike interpretations in an experimental context, then (regardless of the source of these interpretations) they will access non-adultlike interpretations of the linguistic input. Based on the results of the current study, which suggest that children's non-adultlike behavior is due to extragrammatical factors, these non-adultlike interpretations of the input are likely to persist even after children have acquired the adult grammar. This is problematic for acquiring the adult grammar in the first place, then, because it means that the antecedent of PRO is not likely to be a reliable source of evidence for the adult grammar. Children must therefore use a different form of evidence for acquiring the adult grammar, which raises further questions.

First, if children cannot use the elements of a dependency to acquire the restrictions on that dependency, how do they know what can be used as evidence? If this is a problem for adjunct control, then it is likely to be relevant for other

dependencies as well. This point brings up the question of overall frequency. Although structures with adjunct control are not frequent in the input, they are not absent (Gerard in press). However, if children do not use these instances to learn the adjunct control dependency, then the frequency does not contribute to the acquisition problem. If this is true for adjunct control, though, then it is also likely for other types of dependencies, particularly where children access non-adultlike interpretations that would provide evidence for a non-adult grammar.

Next, if children's behavior can still be non-adultlike at age four, what develops so that their responses are eventually adultlike? Depending on the source of the increased accuracy observed for the coloring TVJT, there are various aspects of the deployment system that may become more efficient or faster with development. The current study is not designed to tease these options apart, but they will be investigated in future research.

Conclusion

In this paper we investigated the source of children's non-adultlike behavior for adjunct control, using two different truth value judgment tasks. Children's responses were not as adultlike for a task with higher demands (depending on event ordering) compared to a task with lower demands (depending on the color of an item). This demonstrates that children's interpretations of adjunct control may be affected by extragrammatical factors, which has implications for explaining children's behavior in previous studies, as well as for children's acquisition. Future research will further investigate these extragrammatical factors and how they interact with children's linguistic knowledge.

References

- Adani, Flavia. 2011. Rethinking the acquisition of relative clauses in Italian: towards a grammatically based account. *Journal of Child Language* 38(01), 141–165.
doi:10.1017/S0305000909990250.
- Adler, Allison Nicole. 2006. Syntax and discourse in the acquisition of adjunct control. Massachusetts Institute of Technology. (3 April, 2016).
- Barr, Dale J. 2013. Random effects structure for testing interactions in linear mixed-effects models. *Frontiers in Psychology* 4. doi:10.3389/fpsyg.2013.00328.
- Bates, Douglas, Martin Maechler, Ben Bolker & Steve Walker. 2015. Fitting Linear Mixed-Effects Models Using lme4. *Journal of Statistical Software* 67(1), 1–48.
doi:10.18637/jss.v067.i01.
- Broihier, Kevin & Kenneth Wexler. 1995. Children's acquisition of control in temporal adjuncts. *Papers on Language Processing and Acquisition, MIT Working Papers in Linguistics* 26, 193–220.
- Cairns, Helen Smith, Dana McDaniel, Jennifer Ryan Hsu & Michelle Rapp. 1994. A longitudinal study of principles of control and pronominal reference in child English. *Language* 70(2), 260–288.
- Cairns, Helen Smith, Gloria Schlisselberg, Dava Waltzman & Dana McDaniel. 2006. Development of a Metalinguistic Skill: Judging the Grammaticality of Sentences. *Communication Disorders Quarterly* 27(4). SAGE Publications Inc, 213–220.
doi:10.1177/15257401060270040401.
- Conroy, Anastasia, Eri Takahashi, Jeffrey Lidz & Colin Phillips. 2009. Equal treatment for all antecedents: How children succeed with Principle B. *Linguistic Inquiry* 40(3), 446–486.
- Crain, Stephen & Rosalind Thornton. 1998. *Investigations in universal grammar: A guide to experiments on the acquisition of syntax and semantics*. MIT Press.
- Crain, Stephen, Rosalind Thornton, Carole Boster, Laura Conway, Diane Lillo-Martin & Elaine Woodams. 1996. Quantification without qualification. *Language Acquisition* 5(2), 83–153.
- Davidson, Donald. 1984. *Inquiries into truth and interpretation*, Clarendon. Oxford.
- Dillon, Brian, Alan Mishler, Shayne Sloggett & Colin Phillips. 2013. Contrasting intrusion profiles for agreement and anaphora: Experimental and modeling evidence. *Journal of Memory and Language* 69(2). Elsevier, 85–103.
- Droz, Kenneth. 2001. Children's Weak Interpretation of Universally Quantified Sentences. *Conceptual Development and Language Acquisition*, 340–376.
- Eager, Christopher & Joseph Roy. 2017. Mixed effects models are sometimes terrible. *arXiv preprint arXiv:1701.04858*.

- Engelmann, Felix, Lena A. Jäger & Shravan Vasishth. 2019. The Effect of Prominence and Cue Association on Retrieval Processes: A Computational Account. *Cognitive Science* 43(12), e12800. doi:10.1111/cogs.12800.
- Frazier, Lyn & Janet Dean Fodor. 1978. The sausage machine: A new two-stage parsing model. *Cognition* 6(4), 291–325. doi:10.1016/0010-0277(78)90002-1.
- Fritzley, Heather & Kang Lee. 2003. Do young children always say yes to yes–no questions? A metadevelopmental study of the affirmation bias. *Child Development* 74(5), 1297–1313.
- Fritzley, Heather, Rod C. L. Lindsay & Kang Lee. 2013. Young Children’s Response Tendencies Toward Yes–No Questions Concerning Actions. *Child Development* 84(2), 711–725. doi:10.1111/cdev.12006.
- Gerard, Juliana. in press. Adjunct control and the poverty of the stimulus: availability vs. evidence. In Anne Mucha, Jutta Hartmann & Beata Trawiński (eds.), *Non-Canonical Control in a Cross-linguistic Perspective [Linguistik aktuell/Linguistics Today]*.
- Gerard, Juliana. 2016. The acquisition of adjunct control: Grammar and processing.
- Gerard, Juliana, Jeffrey Lidz, Shalom Zuckerman & Manuela Pinto. 2017. Similarity-Based Interference and the Acquisition of Adjunct Control. *Frontiers in Psychology* 8. doi:10.3389/fpsyg.2017.01822.
- Gerard, Juliana, Jeffrey Lidz, Shalom Zuckerman & Manuela Pinto. 2018. The acquisition of adjunct control is colored by the task. *Glossa: a journal of general linguistics* 3(1). doi:10.5334/gjgl.547.
- German, Diane J. & Rochelle S. Newman. 2004. The impact of lexical factors on children’s word-finding errors. *Journal of speech, language, and hearing research: JSLHR* 47(3), 624–636. doi:10.1044/1092-4388(2004/048).
- Gierut, Judith A. 2001. A model of lexical diffusion in phonological acquisition. *Clinical Linguistics & Phonetics* 15(1–2), 19–22. doi:10.3109/02699200109167624.
- Goodluck, Helen. 1981. Children’s grammar of complement-subject interpretation. In Susan Tavakolian (ed.), *Language acquisition and linguistic theory*, 139–166. Cambridge, MA: The MIT Press.
- Goodluck, Helen. 1998. Children’s interpretation of adjunct PRO: New evidence. *Proceedings of the 22nd Annual Boston University Conference on Language Development*, 293–302. Boston, MA: Cascadia Press.
- Goodluck, Helen. 2001. The Nominal Analysis of Children’s Interpretations of Adjunct Pro Clauses. *Language* 77(3), 494–509. doi:10.1353/lan.2001.0159.
- Goodluck, Helen & Dawn Behne. 1992. Development in control and extraction. In J Weissenborn, Helen Goodluck & Tom Roeper (eds.), *Theoretical issues in language acquisition*, 151–171. Hillsdale, NJ: Lawrence Erlbaum.
- Grice, Herbert P. 1975. Logic and conversation. *Speech acts*, 41–58. Brill.

- Hamburger, Henry & Stephen Crain. 1984. Acquisition of cognitive compiling. *Cognition* 17(2), 85–136.
- Hsu, Jennifer Ryan, Helen Smith Cairns, Sarita Eisenberg & Gloria Schlisselberg. 1989. Control and coreference in early child language. *Journal of Child Language* 16(3), 599–622. doi:10.1017/S0305000900010746.
- Hsu, Jennifer Ryan, Helen Smith Cairns & Robert W. Fiengo. 1985. The development of grammars underlying children’s interpretation of complex sentences. *Cognition* 20(1), 25–48.
- Jäger, Lena A., Lena Benz, Jens Roeser, Brian W. Dillon & Shravan Vasishth. 2015. Teasing apart retrieval and encoding interference in the processing of anaphors. *Frontiers in Psychology* 6. doi:10.3389/fpsyg.2015.00506.
- Jäger, Lena A., Felix Engelmann & Shravan Vasishth. 2017. Similarity-based interference in sentence comprehension: Literature review and Bayesian meta-analysis. *Journal of Memory and Language* 94, 316–339. doi:10.1016/j.jml.2017.01.004.
- Jäger, Lena A., Daniela Mertzen, Julie A. Van Dyke & Shravan Vasishth. 2020. Interference patterns in subject-verb agreement and reflexives revisited: A large-sample study. *Journal of Memory and Language* 111, 104063. doi:10.1016/j.jml.2019.104063.
- Janke, Vikki. 2016. Pragmatic Leads and Null Subjects: When Children Consult Leads and When They Do Not. *Proceedings of the 40th Annual Boston University Conference on Language Development*.
- Janke, Vikki. 2018. Discourse effects on older children’s interpretations of complement control and temporal adjunct control. *Language Acquisition* 25(4), 366–391.
- Janke, Vikki & Laura R. Bailey. 2017. Effects of discourse on control. *Journal of Linguistics* 53(03), 533–565. doi:10.1017/S0022226716000281.
- Key, Alexandra P., Dorita Jones & Sarika U. Peters. 2016. Response to own name in children: ERP study of auditory social information processing. *Biological Psychology* 119, 210–215. doi:10.1016/j.biopsycho.2016.07.016.
- Kush, Dave & Brian Dillon. 2020. *Principle B constrains the processing of cataphora: Evidence for syntactic and discourse predictions*. Preprint. Open Science Framework. doi:10.31219/osf.io/mxtc6.
- Landau, Idan. 2021. *A Selectional Theory of Adjunct Control*. MIT Press.
- Laurinavichyute, Anna, Lena A. Jäger, Yulia Akinina, Jennifer Roß & Olga Dragoy. 2017. Retrieval and Encoding Interference: Cross-Linguistic Evidence from Anaphor Processing. *Frontiers in Psychology* 8. Frontiers. doi:10.3389/fpsyg.2017.00965.
- Lewis, Richard L. & Shravan Vasishth. 2005. An activation-based model of sentence processing as skilled memory retrieval. *Cognitive science* 29(3), 375–419.

- Lust, Barbara, Larry Solan, Suzanne Flynn, Catherine Cross & Elaine Schuetz. 1986. A Comparison of Null and Pronoun Anaphora in First Language Acquisition. In Barbara Lust (ed.), *Studies in the Acquisition of Anaphora: Defining the Constraints* (Studies in Theoretical Psycholinguistics), 245–277. Dordrecht: Springer Netherlands. doi:10.1007/978-94-009-4548-7_7.
- McDaniel, Dana & Helen Smith Cairns. 1990a. The processing and acquisition of control structures by young children. *Language processing and language acquisition*, 313–325. Springer. (3 April, 2016).
- McDaniel, Dana & Helen Smith Cairns. 1990b. The child as informant: Eliciting linguistic intuitions from young children. *Journal of Psycholinguistic Research* 19(5), 331–344.
- McDaniel, Dana, Helen Smith Cairns & Jennifer Ryan Hsu. 1991. Control principles in the grammars of young children. *Language Acquisition* 1(4), 297–335.
- Omaki, Akira & Jeffrey Lidz. 2015. Linking parser development to acquisition of syntactic knowledge. *Language Acquisition* 22(2), 158–192.
- Özge, Duygu, Aylin Küntay & Jesse Snedeker. 2019. Why wait for the verb? Turkish speaking children use case markers for incremental language comprehension. *Cognition* 183, 152–180. doi:10.1016/j.cognition.2018.10.026.
- Parker, Dan, Sol Lago & Colin Phillips. 2015. Interference in the processing of adjunct control. *Frontiers in psychology* 6.
- Pinto, Manuela & Shalom Zuckerman. 2018. Coloring Book: A new method for testing language comprehension. *Behavior Research Methods*. doi:10.3758/s13428-018-1114-8.
- R Core Team. 2015. *R: A language and environment for statistical computing*. R Foundation for Statistical Computing (R Foundation for Statistical Computing). Vienna, Austria.
- Storkel, Holly L. & Michele L. Morrisette. 2002. The Lexicon and Phonology: Interactions in Language Acquisition. *Language, Speech, and Hearing Services in Schools* 33(1), 24–37. doi:10.1044/0161-1461(2002/003).
- Syrett, Kristen & Jeffrey Lidz. 2011. Competence, performance, and the locality of quantifier raising: Evidence from 4-year-old children. *Linguistic Inquiry* 42(2), 305–337.
- Troia, Gary A., Froma P. Roth & Grace H. Yeni-Komshian. 2018. Word Frequency and Age Effects in Normally Developing Children’s Phonological Processing. Research-article. *ASHA Wire*. American Speech-Language-Hearing Association. World. doi:10.1044/jshr.3905.1099.
- Villata, Sandra, Whitney Tabor & Julie Franck. 2018. Encoding and Retrieval Interference in Sentence Comprehension: Evidence from Agreement. *Frontiers in Psychology* 9. doi:10.3389/fpsyg.2018.00002.
- Wexler, Kenneth. 1992. Some issues in the growth of control. In R Larson, S Iatridou, U Lahiri & J Higginbotham (eds.), *Control and grammar*, 253–295. Springer.