Adjunct control and the poverty of the stimulus: availability vs. evidence

Juliana Gerard
Ulster University

Subject control in non-finite adjuncts is observed across languages (as in 'John called Mary before running to the store'). Research on the acquisition of adjunct control has generally focused on the relevant grammatical components and when they are acquired. This paper considers these components in the context of the linguistic input to ask how control in adjuncts is acquired.

Although adjunct control is available in the input, the instances themselves do not provide evidence for abstract syntactic relations. Implications are considered for linguistic dependencies and the evidence in the input.

1. Introduction

This paper focuses on obligatory control in non-finite adjuncts, as in (1):

(1) John$_1$ called Mary$_2$ after PRO$_{1/*2/*3}$ drawing a picture.

In particular, adjunct control is used as a case study for the role of the linguistic input in acquiring dependencies: while some properties of adjunct control are observed across languages, others are language-specific. Additionally, exceptions to canonical control structures raise questions about the type of information needed from the input.

In (1), the adjunct subject PRO is obligatorily controlled by the main clause subject *John*. This pattern is observed across languages, and is captured by high attachment of the adjunct clause and c-command by the

controller (Chomsky 1981).[1] Therefore, evidence for these features must be available in the linguistic input or they must be innate (Chomsky 1965).

The goal of this paper is to evaluate these features and their predictions for the input. For example, if evidence is available for attachment height and the c-command dependency, it may be available directly or by generalization from other structures. Meanwhile, if these features are innate (specified in Universal Grammar, or UG), then evidence is needed for language-specific aspects of the dependency.

Obligatory subject control is relatively consistent for temporal adjuncts like (1), although some variation is observed both within and across languages, e.g. in contexts with a logophoric or arbitrary controller (Williams 1992; Landau 2015, 2017; Green 2018a). Thus, a complete account of adjunct control should consider:

i. how to acquire abstract properties like attachment height and c-command
(where the abstract properties are not observed directly)
ii. how to identify the different types of adjunct control, given the evidence in the input
iii. implications of innate properties (i.e. not acquired from the input) for a model of language

This paper will focus primarily on (i), concluding with a discussion of (ii) and (iii).

2. How to acquire control

The primary question of this paper is how control in adjuncts is acquired. The following hypotheses are considered:

[1] This paper is based on these components, but may also be considered in the context of other frameworks; importantly, adjunct control involves a locality constraint which is structurally defined. This constraint is the focus of this paper.

(a) evidence for attachment height and c-command is available in the input, either
    i. by observing instances of adjunct control directly or
    ii. by generalizing the relevant features from similar structures.
(b) evidence for these features is not in the input, and the features are specified in UG.

The following sections consider each of these options in turn. The analysis, based on a critical review of previous behavioral data and on relevant corpus data, finds that evidence for attachment height and c-command is not likely to be available in the linguistic input. This includes both direct evidence (from observing instances of adjunct control in the input) and indirect evidence (by generalizing from similar structures). If attachment height and c-command are innate, however, this raises further questions about the role of the linguistic input, and makes predictions about children's behavior in experimental contexts (to be tested in future research).

2.1 *A note on domain specificity*

The learning strategies discussed below may be domain-specific or more domain-general. For example, strategies that depend on a syntactic category or transformation must involve some domain specificity in order to focus on a linguistic feature. However, this contrast is tangential to the primary aim. Rather, the focus of this paper is on the role of the input: are the grammatical components of adjunct control acquired or innate? The following sections consider the options.

3. From the input

If either attachment height or c-command by the controller are acquired from the linguistic input, then explicit predictions are made about the evidence in the input. Two types of evidence will be considered here: first, the conditions are spelled out for inferring the correct attachment height or c-command by observing instances of adjunct control directly. Next, these

features may be generalized to adjunct control from similar structures, which may be more frequent or salient in the input.

3.1 *Direct observation*

For attachment height and c-command to be inferred by observing instances of adjunct control, there must be instances of adjunct control available in the input. However, while adjunct control is necessary, it is not sufficient for providing evidence for the relevant abstract features; other factors to consider include the prerequisite linguistic knowledge and children's perception of the input. These additional factors are discussed in the following sections, with the assumption that adjunct control is indeed available in the input.

3.1.1 Attachment height

If children need evidence for adjunct attachment height, then incorrect attachment is predicted before the relevant evidence is encountered in the input, which would result in non-adultlike interpretations of adjunct PRO. Indeed, children in previous studies have accepted a range interpretations, and one prominent account is misattachment of the adjunct to the main clause (Goodluck 1981; Hsu, Cairns, and Fiengo 1985; McDaniel, Cairns, and Hsu 1991; Cairns et al. 1994; Adler 2006). Two primary forms of evidence have been considered in previous studies, which make different assumptions about children's pre-existing knowledge.

3.1.1.1 Lexical learning (Cairns et al. 1994)

To account for children's behavior, Cairns et al (1994) propose different non-adult grammar types, which predict non-adultlike interpretations before children acquire the adult grammar. These grammar types involve high attachment of the adjunct to the main clause (coordination) or low attachment (with c-command by the main clause object). Here, an important distinction is made between types of accounts: these non-adult grammar types can explain children's *behavior* in the study; however, the grammar types alone do not provide an account of *acquisition* - i.e. how a learner can transition from a non-adult grammar to the adult grammar.

To account for children's acquisition, Cairns et al (1994) cite the Lexical Learning Hypothesis (Wexler and Chien 1985), noting that children must link each complementizer form with its selectional properties. They suggest that incorrect attachment results from mapping a complementizer form first to a coordinating structure, before acquiring the correct mapping for a non-finite adjunct. Evidence for the correct attachment would therefore be available with any instance of a given complementizer (not just as a non-finite adjunct), with the transition to the adult grammar resulting from "accretion of lexical and semantic knowledge" for each complementizer (Cairns et al. 1994:264).

This description accounts for the transition to the adult grammar; however, it does not involve the acquisition of syntactic structure. It assumes instead that children already have the relevant abstract knowledge of coordination and subordination, with incorrect form-structure mappings. That is, it is not explained how the correct attachment height is acquired in general for adjuncts, as the structure that complementizers are mapped to. If adjunct attachment height is assumed as preexisting knowledge, then it must be innate. Alternatively, another form of evidence may be available in the input for attachment height.

3.1.1.2 Adjunct misanalysis (Adler 2006)

In a different misattachment account, Adler (2006) suggests that the syntactic contrasts between non-finite adjuncts and coordinated clauses may be used as cues to attachment height. For example, the verb form in non-finite adjuncts contrasts with the finite form in coordinated clauses:

(2) a. John eats cake before opening/*opens presents.
b. John eats cake and (then) opens/*opening presents.
adapted from Adler (2006)

Other contrasts involve transformations; for example, cleft structures are possible with adjuncts but not coordinate clauses:

(3) a. It was before opening presents that Mary cut the cake.
b. *It was and John opened presents that Mary cut the cake.

Similarly, different profiles are observed for extraction:

(4) a. What$_i$ did you eat t_i before John opened presents?
b. *What$_i$ did you eat t_i and (then) John open presents?

Importantly, these examples involve positive evidence (Berwick 1985): in (2) the contrast in verb form (or finiteness) is a cue to the contrast in clause type, while in (3) and (4), the transformation itself is a cue - since the sentences are not possible with a coordinated clause, any instances in the input would need to be represented with an adjunct clause (Adler 2006).

However, the above evidence is still problematic for learning attachment height. In (2), the contrast in verb form aligns with the contrast in attachment height: that is, coordinated clauses and non-finite clauses have different verb forms and different attachment heights. This strategy makes the wrong predictions for finite adjuncts, though, which also have a finite verb form (grouping finite adjuncts with coordinated clauses):

(5) John eats cake before he opens/*opening presents.

This miscategorization may be avoided if the contrast in (2) is applied to a subset of the input data. However, this would involve domain-specific knowledge about which data to use for learning, merely shifting the learning problem rather than addressing it.

Meanwhile, the sentences in (3) and (4) must be represented accurately in order to be used as evidence for the correct attachment height. However, the influence of an immature parser, along with high sentence processing costs may affect the reliability of this evidence.

More broadly, both types of evidence discussed by Adler (2006) rely on prior knowledge of a contrast in attachment height between adjuncts and coordinated structures. Moreover, similar to the approach by Cairns et al (1994), the relevant learning strategies involve mapping a lexical item (complementizer) to abstract structure (adjunct clause), by abandoning an initial incorrect mapping (coordinated clause). These mappings are important, but they require the attachment height for adjuncts to have already been acquired. Again, attachment height must either be innate here, or acquired using another form of evidence. A final possibility for attachment height is discussed in the following section.

3.1.1.3 Binding across clauses

The next type of evidence to consider for attachment height involves binding relations across clauses, as in (6) and (7):

(6) He_1 called Mary before $John_{*1/2}$ left for the store.
(7) John called her_1 before PRO meeting $Mary_1$ at the store.

In (6), the pronoun *he* c-commands *John*, and co-reference is ruled out by Principle C (Chomsky 1981). However, co-reference is possible if the adjunct is attached high. Thus, if children have a grammar with high attachment, negative evidence is needed against co-reference in sentences like (6), which may then be used to infer the correct (lower) attachment height.[2]

Meanwhile, syntactic evidence against a low attaching adjunct is seen in sentences like (7), with co-reference between *her* and *Mary*. If children have a grammar with low attachment, then coreference in the input with this configuration would provide positive evidence for the correct (higher) attachment height.

For both (6) and (7), the relevant evidence involves several assumptions which are problematic for acquisition. First, evidence against the coreference in (6) might be available in the form of indirect negative evidence (Xu and Tenenbaum 2007); however, previous research on children's acquisition of Principle C finds that children already reject coreference in this configuration from as young as 3 years of age (Crain and McKee 1985; Crain and Thornton 1998; for reviews, see Lust, Eisele, and Mazuka 1992; Guasti 2017). This timeline is inconsistent with studies on adjunct control, where children's non-adultlike interpretations were observed until 5-6 years of age.

Alternatively, children might acquire a high attachment grammar initially but get evidence for the adult grammar before age 3. However, if the relevant evidence involves referential dependencies across multiple

[2] As (6) is finite, this strategy involves an additional generalization from finite to non-finite adjuncts (discussed further below).

clauses, the timeframe is further limited by children's parsing abilities at this age.

More importantly, using binding across clauses as evidence for attachment height involves the crucial assumption that the relevant configurations will be available in the linguistic input. However, for both (6) and (7), the critical anaphoric relations are highly infrequent, especially if the relevant timeframe is limited by other factors like the developing parser (Sutton 2015; Gerard 2016). Furthermore, this type of evidence depends on the coreference interpretation, which children may not always access: if a different referent is assigned in children's perception of the input, i.e. the linguistic *intake*, then this will provide evidence for the incorrect attachment height (Lidz and Gagliardi 2015; Omaki and Lidz 2015). Thus, it is unlikely that binding relations alone are used as evidence for attachment height for non-finite adjuncts.

Attachment height will be addressed again in the section on generalization; the following section considers the evidence for inferring a c-commanding controller.

3.1.2 C-command by the controller

Inferring the c-command relation between the main clause subject and adjunct PRO is a two-step process:

1. Identify the set of possible antecedents for adjunct PRO (i.e. the main clause subject).
2. Determine that the dependency is due to c-command, as opposed to e.g. a discourse or agent preference or based on a property like animacy, which are also likely to involve the main clause subject

It is assumed for these steps, particularly step 2, that the adjunct is attached correctly - either by having acquired the feature already, or if adjunct attachment is specified in UG (Goodluck and Behne 1992).

Meanwhile, these steps must be indirect on some level: with just a single instance of adjunct control in the input, the interpretation of PRO is consistent with multiple grammars. For example, in addition to a strict subject (adult) grammar, the coreference in (1), repeated below as (8), is

also consistent with an agent grammar, a sentence-internal grammar, a free reference grammar (where *John* is retrieved from the discourse), and others.

(8) John$_1$ called Mary$_2$ after PRO$_{1/*2/*3}$ drawing a picture.

All things equal, inferring that the antecedent of PRO is the main clause subject therefore requires multiple instances of adjunct control. However, children's interpretations in previous studies suggest that this inference will be problematic, for any type of learning mechanism (domain-specific or domain-general).

Consider the range of interpretations observed in previous studies: in addition to a strict subject (adultlike) interpretation, children's responses were also categorized as strict object control, sentence internal, and free reference. While children with a strict object control grammar would have given only non-adultlike responses, the set of interpretations accepted by children with a sentence internal or free reference grammar are a superset of the interpretations allowed by the strict subject grammar. These relations are illustrated in Figure 1. Importantly, these non-adultlike grammars can account for children's behavior, but they introduce complications for an account of acquisition which depends on the antecedent of PRO.

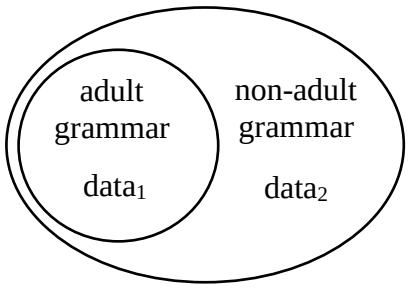


Figure 1. Subset-superset relation between the adult grammar (strict subject) and non-adult grammars for adjunct PRO (sentence-internal and free reference). While the adult grammar includes data$_1$ but not data$_2$, the non-adult grammar includes both data$_1$ and data$_2$.

Traditionally, children with a non-adult grammar will encounter some form in the input which is consistent with the adult grammar but not with the non-adult grammar, and this form will be evidence for the adult grammar (Gold 1967; Grimshaw and Pinker 1989; Pinker 1979, 2009). This

logic is discussed in section 3.1.1.2 above for encountering syntactic evidence against a coordination grammar. However, when the interpretations in the adult grammar are a subset of the interpretations in the non-adult grammar, negative evidence is needed to discard the larger grammar in favor of the smaller one (Berwick 1985; Gold 1967; Baker 1979; Manzini and Wexler 1987; Pinker 2013; Heinz and Riggle 2011).

Since Gold's (1967) proposal, different systems have been considered to handle this type of transition. For example, the Subset Principle (Berwick 1985) allows for learning from positive evidence if there is a specific acquisition procedure that always chooses a subset grammar when the evidence is consistent with multiple grammars. Meanwhile, under the size principle, smaller hypotheses are considered to be more likely than larger hypotheses (which generate a superset of the data generated by a smaller hypothesis), and exponentially more likely as more data that is observed that is compatible with both hypotheses (Tenenbaum 1999; Tenenbaum and Griffiths 2001; Xu and Tenenbaum 2007). However, this logic does not work with evidence for the subject as the antecedent of PRO, and highlights a more general problem with acquiring syntactic constraints on anaphora.

A non-adult grammar which allows a superset of the interpretations in the adult grammar is represented in Figure 1 - for example, a free reference grammar, which allows the subject control reflected in (8), but also the object control interpretation and an external interpretation. The subset grammar is the strict subject (adult) grammar, which allows only a subject control interpretation. Under the size principle, children should transition from the superset grammar to the subset grammar by observing instances of adjunct control in the input with a subject control interpretation, represented by $data_1$ in Figure 1. The subset grammar should be considered to be more likely if data like $data_1$ occur in the input, which should indeed be the only type of data in the input (speech errors notwithstanding). However, this overlooks the contrast between the input and the intake, and the finding from previous studies that children allowed non-adultlike interpretations of adjunct PRO. If these children's grammars were not adultlike, then they would also allow non-adultlike interpretations of the input, represented by $data_2$ in Figure 1. Crucially, $data_2$ will provide evidence against the adult grammar *and* for the non-adult grammar.

As a result, children's interpretations of adjunct PRO are not a reliable cue for inferring the c-command relation. Moreover, other syntactic dependencies face a similar dilemma: if children accept a wider range of interpretations in an experimental context, then the same interpretations are likely to be available in the linguistic input. Further implications are discussed in the final sections.

If the grammatical components of adjunct control are not inferred directly - from instances of adjunct control in the input - then evidence may instead be available from other structures, which may be generalized to structures with adjunct control.

3.2 *Generalization from similar structures*

The following sections will consider the possibility of generalizing attachment height and c-command to sentences with adjunct control from two similar structures: complement control, where the dependent element has the same form; and finite adjuncts, with a similar syntactic context.

3.2.1 Complement control

In sentences with complement control (as in (9), below), the same c-command relation is generally observed for the controller - that is, the closest c-commanding NP - with the same (null) form of PRO:

(9) a. $John_1$ wanted PRO_1 to run to the store.
b. $John_1$ told $Mary_2$ $PRO_{*1/2}$ to run to the store.

In previous studies, children have exhibited adultlike behavior for complement control before adjunct control (Hsu et al. 1985; McDaniel et al. 1991; Cairns et al. 1994); however, children still accepted a wider range of interpretations initially, albeit at a younger age than for adjunct control. This suggests that children do not infer the antecedent of PRO from sentences with complement control, since the non-adultlike interpretations would provide incorrect evidence in the input in the same way as discussed above for adjunct control.

A generalization strategy also makes several assumptions: first, if children did infer the antecedent for complement control, then the same

inference must not also be made for adjunct control. Next, if children generalize from complement control to adjunct control, this assumes that the relevant generalization is not made in the reverse direction, from adjunct control to complement control. Finally, adjunct control and complement control share various features; if children do generalize the correct features, then they must avoid generalizing others (e.g. attachment height or verb form).

These arbitrary assumptions about what is generalized suggest that children do not generalize from complement control to adjunct control, at least for a property like the antecedent of PRO.

3.2.2 Finite adjuncts

For the purposes of identifying the controller, finite adjuncts have the same attachment height as non-finite adjuncts, as demonstrated by the co-reference in (10) between *her* and *Mary*:

(10) John_1 called her_2 before he_1 met Mary_2 at the store.

Therefore, if children could acquire the attachment height for finite adjuncts from the linguistic input, then this might then be generalized to non-finite adjuncts.

However, the evidence needed for attachment height with finite adjuncts has the same problems discussed above for non-finite adjuncts - for example, evidence in the form of binding relations across clauses is unlikely to occur in the input, falling short of explaining how attachment height is acquired in general.

Additionally, the same assumptions are made for finite adjuncts as the ones outlined above for complement control: if children did infer attachment height for finite adjuncts, then the same inference must not also be made for non-finite adjuncts. Next, if children did generalize from finite adjuncts to non-finite adjuncts, this assumes that the relevant generalization is not made in the reverse direction, from non-finite adjuncts to finite adjuncts. Finally, finite adjuncts and non-finite adjuncts share various features; if children do generalize the correct features, then they must avoid generalizing other ones (e.g. antecedent or verb form).

Like for complement control, these arbitrary assumptions suggest that children do not generalize a feature like attachment height from finite adjuncts to non-finite adjuncts. However, the final point about the incorrect features makes predictions about the linguistic input which can be tested. For example, if children generalize the antecedent from complement control, then adultlike patterns should be observed for both frames, which does not distinguish between generalization and other possible strategies; however, if children generalize the antecedent from finite adjuncts - which allow free reference, rather than strict subject control - then a similar pattern should also be observed for children's productions of non-finite adjuncts. This prediction is tested in the following section.

3.2.3 Testing input predictions: a corpus analysis

The linguistic input is represented here by transcripts of speech to children from CHILDES (MacWhinney 2000). The corpus - all of the CHILDES transcripts in North American English - contained roughly 2.8 million utterances and nearly 10 million words, which represents around 3 years of speech to children (Akhtar et al. 2004; Hart and Risley 1995). While the children in the transcripts ranged in age from less than 12 months to 12 years, 71% of the transcripts focused on children from 2-4 years of age - precisely the age range of interest for acquiring adjunct control, given the developmental trajectory observed in previous studies (specifically, 29% of transcripts were from 2 year olds, 28% from 3 year olds, and 14% from 4 year olds).

For testing predictions about generalization, the speech to children in the transcripts represents the input, with the caveat that the input does not perfectly reflect the intake, as discussed above. However, it can provide an informed estimate of children's exposure to structures like adjunct control. Similarly, children's own speech is an observable behavior, representing their linguistic competence filtered through a performance system. However, children's productions can provide an estimate of their competence for structures like adjunct control. Inferences can also be made about the prerequisite knowledge for a particular learning strategy - for example, interpreting binding relations across clauses may be more likely once children start to produce multi-clause utterances.

3.2.3.1 Predicted distributions

If non-finite adjuncts occur in the input, they should occur only with subject control interpretations (since adults should not produce any ungrammatical instances of adjunct control). In contrast, finite adjuncts allow free reference, as in (11):

(11) $John_1$ called $Bill_2$ after $he_{1/2/3}$ drew a picture.

Therefore, finite adjuncts may have the same subject control distribution as non-finite adjuncts *or* they may have a higher frequency of other internal antecedents (*Bill* in (11)) or external antecedents. If children do generalize from finite to non-finite adjuncts, these possibilities for finite adjuncts make different predictions:

(a) If the distribution of finite adjunct subjects in the input is the same as the distribution of non-finite adjuncts (i.e. mostly subject antecedents), then children should also have a subject bias for both finite and non-finite adjuncts. This may be expected given the general bias for a subject antecedent in sentences like (11), all things equal (Arnold 2001; Kehler and Rohde 2013).

(b) If the distribution of finite adjunct subjects in the input is different from non-finite adjuncts - e.g. with a higher proportion of other internal antecedents or a higher proportion of external antecedents, then children should generalize this free reference pattern to non-finite adjuncts. This predicts that children's productions of non-finite adjuncts will include not just subject control, but other antecedents as well. This non-adultlike pattern would be consistent with children's behavior in previous studies, where children accepted other antecedents for non-finite adjuncts. However, further evidence would then be needed for non-finite adjuncts for strict subject control, which must not be generalized back to finite adjuncts.

Meanwhile, if children do not generalize from finite to non-finite adjuncts, then the distribution of finite adjunct antecedents should not predict children's distribution for non-finite adjuncts.

3.2.3.2 Search methods

Instances of adjunct control were identified by searching for each complementizer followed by the string "ing" (Broihier and Wexler 1995). Non-finite complementizers included in the search were *after, before, while, without,* and *instead of.* Finite adjuncts were identified using the parsed mor tier by searching for each complementizer followed a pronoun, a bare noun, or determiner. Finite complementizers included in the search were *after, before,* and *while*. Note that *without* and *instead of* are only possible in a non-finite frame in English, and were not included in the search for finite frames.

These searches returned a number of false positives (e.g. "what happens after spring"), so the search results were hand coded for actual instances of non-finite or finite complementizers. The utterances were also hand coded for the antecedent of the adjunct subject as either the main clause subject, direct object, other internal referent, or external referent. For example, in (12), the subject of the finite adjunct *she* co-refers with the main clause subject *Mommy,* so this utterance was coded as a main clause subject interpretation:

(12) Mommy fixed it before she left.

Examples of each type of interpretation are given in Table 1.

Finiteness	Antecedent	Utterance
Non-finite	Subject	I fell asleep after counting to three
Non-finite	Subject	I want a story before going to sleep
Finite	Subject	I can put this back on after I take it off
Finite	Object	We just caught the fish before it went down the drain
Finite	Other internal	You can't sit in her chair while she's there
Finite	External	You can keep making cookies while I load up the car

Table 1. Examples of non-finite and finite utterances from CHILDES

3.2.3.3 Corpus analysis results

The results of the corpus analysis are presented in Figure 2 (for adults) and Figure 3 (for children), by finiteness and complementizer. The counts are raw frequencies from the full corpus, and counts for the main clause object are included in the "other internal" condition.

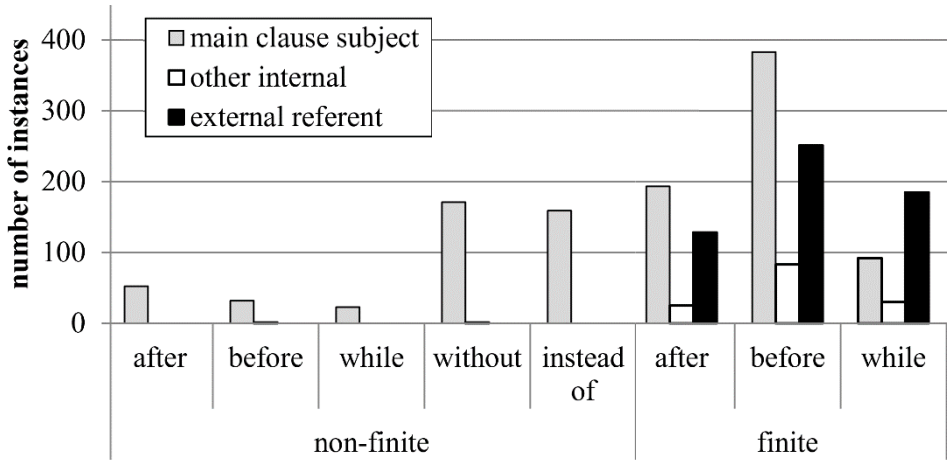


Figure 2. Frequency of antecedents in non-finite and finite adjuncts in CHILDES, speech to children (linguistic input).

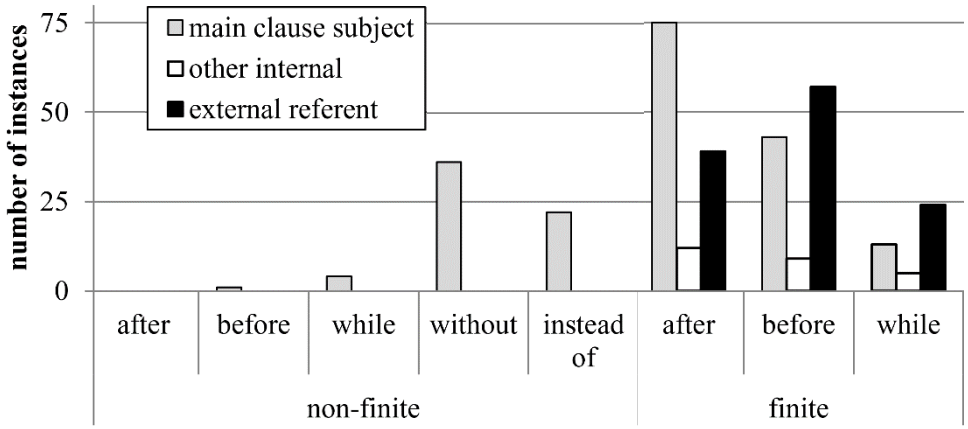


Figure 3. Frequency of antecedents in non-finite and finite adjuncts in CHILDES, children's speech (note the scaled y axis).

Non-finite and finite adjuncts were produced by both children and adults, with children producing 14% of the instances overall (48 out of 345 instances). Adjunct control appeared in speech to children from the earliest transcript - particularly with the complementizers *without* and *instead of* - and continued throughout. In children's own speech, adjunct control is observed as early as age 2, and becomes more frequent with age.

As predicted for adults, only subject control was observed for the non-finite adjuncts (in Figure 2, non-finite *after, before, while without,* and *instead of*). Meanwhile, a subject bias was not observed for finite adjuncts; rather, there was a roughly even split for the antecedents of finite adjuncts between the main clause subject and an external antecedent (in Figure 2, finite *after, before,* and *while*). Therefore, if children generalize the antecedent from finite adjuncts to non-finite adjuncts, external antecedents are also predicted for non-finite adjuncts.

Indeed for *finite* adjuncts, the distribution of antecedents for children was nearly identical to the distribution of finite adjuncts for adults (in Figure 3, finite *after, before,* and *while*), showing that children are sensitive to the distribution of antecedents in the input. However, just like adults, children produced only subject control for non-finite adjuncts (in Figure 3, non-finite *after, before, while, without,* and *instead of*). This result has suggests that:

- Children do not generalize the antecedent of the subject from finite adjuncts to non-finite adjuncts.
- Children are sensitive to finiteness in adjuncts from the start, with no stage where children treat non-finite adjuncts like finite adjuncts.

Another contrast to note in Figures 2 and 3 is the frequency of non-finite adjuncts with *without* and *instead* of - which can only appear in a finite frame - compared to the frequencies for non-finite *after, before,* and *while*. The instances with *without* and *instead of* are much more frequent than with *after, before,* and *while;* additionally, for *after, before,* and *while*, the main clause subject interpretation is much more frequent in the finite frame than in the non-finite frame, for both children and adults. This suggests that both children and adults will have a preference for the finite frame when both are possible, and consequentially are likely to have a general expectation for the finite frame over the non-finite frame; however, children are still aware of the optionality.

Based on these frequencies, one final option is that children generalize from the more frequent non-finite adjuncts - *without* and *instead of* - to the less frequent ones - *after, before,* and *while*. This is possible, but it does not explain how adjunct control is acquired in general, as a syntactic dependency (or account for the more frequent non-finite adjuncts).

In sum, the above data suggest that children do not receive misleading evidence in the input, and their productions are consistent with adultlike competence. No stage is observed where children's productions of adjunct control are not adultlike, consistent with the discussion in the previous sections: evidence is not likely to be available in the input for attachment height or c-command by the controller, either directly or by generalizing from other structures.

3.2.3.4 A note on dialectal variation

The above corpus analysis reports data from children acquiring North American English. While this includes various regional dialects, it excludes other dialects of English which may have a different developmental trajectory. To control for potential dialectal variation, a secondary analysis

examined non-finite adjuncts in UK English. In general, the results were comparable to North American English:

- Instances of adjunct control in the input were observed on average at the same ages in UK English as in in North American English (Figure 4). While some instances were observed earlier in North American English, this may also be due to the general distribution of the transcripts. Importantly, the distributions between the critical ages of 2-4 years were essentially even.
- The same higher frequencies for adjuncts with *without* and *instead of* were also observed in the UK English corpus, compared to *after*, *before*, and *while* (Figure 5).

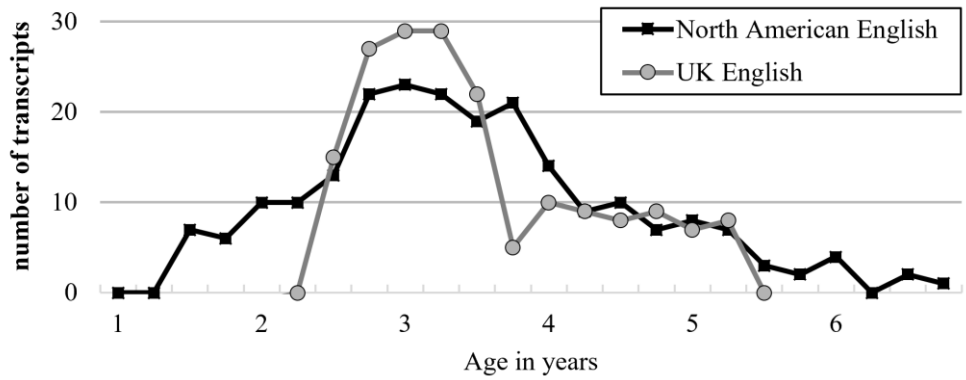


Figure 4. Number of transcripts in CHILDES with adjunct control, by age and corpora: North American English and UK English

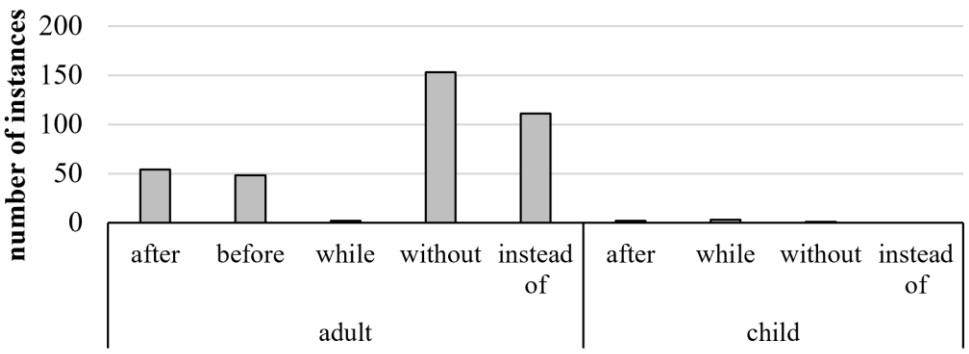


Figure 5. Instances of adjunct control in the UK English CHILDES corpus, by adults and by children

Aside from the lower frequency of *while*, the main difference between the North American and UK English data is that children's productions accounted for far fewer of the instances of adjunct control at 2% of the instances of the UK corpus (7 out of 390 instances) compared to 14% in the North American corpus. If this contrast in production is indicative of children's acquisition, then this has more general implications for the role of the linguistic input, given the similar input distributions for adjunct control in UK and North American English.

Importantly for both corpora, adjunct control is available in the input; however, as discussed in the previous sections, this is not sufficient to provide evidence for the abstract syntactic properties. These properties of adjunct control must then be innate, i.e. part of UG.

4. Universal Grammar

Even though adjunct control itself is available in the input, the relevant evidence is not - at least for acquiring the main syntactic components of adjunct control, attachment height and c-command by the controller. This suggests that these properties are part of UG, which has implications for the hypothesis space of possible grammars considered by a learner. In particular, a learner will only consider the grammars where these properties are adultlike.

If evidence for attachment height and c-command is not in the input, this raises the question of what *is* in the input. What features of adjunct control must be acquired? Predictions are also made for children's acquisition which may be tested empirically.

4.1 *Role of the input*

If the properties of adjunct control are abstract universals, then the input is needed for any variation. For example, finiteness distinguishes non-finite adjuncts from finite adjuncts and conjoined clauses. If tense can be used as a cue for the type of dependency, then it may be one of the features to acquire from the input for adjunct control.

4.1.1 Finiteness

Compared to the abstract syntactic properties, morphological tense is more accessible in the input: the contrast between finite and non-finite verbs is generally realized overtly, and is not limited to adjunct control. For example, the contrast between finite and non-finite clauses is also relevant for complement control, as well as syntactic bootstrapping for verb learning (Harrigan, Hacquard, and Lidz 2019). Indeed, in the corpus data discussed in the previous section, children's production reflected an early sensitivity to finiteness.

An additional cue to adjunct control is the form of the subject - while finite adjuncts generally have an overt subject, in non-finite adjuncts the subject is not pronounced (from the point of view of the learner). Therefore, a learner may look for an empty subject *or* non-finite morphology to identify an adjunct control dependency. Of course, this raises an additional question: would these cues be weighted differently in a language depending on their availability or reliability (Kempe and MacWhinney 1999)? For example, for languages which allow the subject to be dropped (e.g. pro drop, topic drop), the empty subject would not be as helpful for identifying an adjunct control dependency, since finite verbs may also appear without a subject (Haegeman 2000; Holmberg, Nayudu, and Sheehan 2009; Huang 1984; Sundaresan 2014; Nunes 2014; Wu 1992). However, the probability of an empty subject is much higher in a non-finite clause than in a finite clause, even for languages which allow subject drop (since the probability of an overt subject in a non-finite clause is essentially zero). Children are sensitive to these contrasts in probability (for a review see Newport 2016). Therefore, if children use tense or subject form as a cue for adjunct control, then cross-linguistic predictions may be made for acquisition based on (a) the availability of tense (for languages which express tense overtly vs covertly), and (b) the reliability for predicting an empty subject in non-finite vs finite verbs.

For example, the cue to retrieve an antecedent is the missing subject, but if a missing subject may occur in a finite or non-finite clause (as in languages which allow the subject to be dropped), then tense information is also needed to identify the grammatical antecedent. Meanwhile, in languages which do not allow subject drop, if empty subjects are associated with non-finite clauses then an antecedent may be identified without tense information. If the retrieval mechanism is deployed as soon as possible, then

children's parsing strategies may vary depending on these cues (to be tested in future research).

4.1.2 Complementizers

Another feature of adjunct control which varies cross-linguistically is the specific complementizers and the clauses that they select. For example, *without* may appear in a finite frame in both German and Dutch, but not in English:

Non-finite:

(13) a. John cooks without PRO sleeping
b. Fritz kocht ohne PRO zu schlafen
Fritz cooks without PRO to sleep
"Fritz cooks without sleeping"
adapted from Ller (1995)
c. Hij gaf, zonder PRO_1 het te weten,
He gave, without PRO it to know
het juiste antwoord
the right answer
"He gave, without knowing it, the right answer."
adapted from dutchgrammar.com

Finite:

(14) a. *John cooks without that he sleeps
b. Fritz kocht ohne dass er schläft
Fritz cooks without that he sleeps
"Fritz cooks without 'that he sleeps'"
c. Hij gaf, zonder dat hij het wist,
He gave, without that he it knew
het juiste antwoord
the right answer
"He gave, without 'that he knew it,' the right answer."

Therefore, children must learn the form for each complementizer, and whether it selects a finite clause, non-finite, or both. Alternatively, some complementizers may be categorized based on a particular feature to be

learned in groups, although that would introduce the additional question of how this feature is acquired.

The issue of adjunct complementizers is relevant for any acquisition account of adjunct control: complementizers must be distinguished from conjoined clauses and complement clauses. If attachment height is an expected (innate) contrast, then the learning problem will involve identifying the complementizer forms and their selected clauses, and other lexical and semantic properties as discussed by Cairns et al (1994). This has implications, then, for children's competence and the expected developmental trajectory. These are discussed further in the following sections.

4.2 *Competence and acquisition*

In previous studies on the acquisition of adjunct control, children's behavior has generally been attributed to a non-adultlike grammar. However, if both attachment height and c-command by the controller are already part of UG, then these properties of adjunct control would not need to be acquired from the input. Instead, the input would be used for mapping overt forms (like tense and complementizers) to the abstract structure in UG. This predicts that children might sometimes make the wrong mappings, but no stage should be observed with non-adultlike attachment height or a non-adultlike controller.

As seen in the corpus analysis above, this prediction is borne out in children's production; however, it does not explain the non-adultlike behavior observed for children's comprehension in previous studies. If children's competence was adultlike, why would they access non-adultlike interpretations? As discussed in the previous sections, children's interpretations reflect their linguistic competence filtered through an immature parser. That is, children may have acquired the adult grammar, but processing limitations may interfere with the deployment of this grammatical knowledge in an experimental setting. These processing limitations may involve parsing mechanisms for antecedent retrieval (Gerard et al. 2017), as well as the complexity of the task itself (Gerard et al. 2018; Gerard submitted).

This brings back the issue of domain specificity: for children to access adultlike interpretations consistently, what is developing and is it

specific to language? For example, domain-general memory mechanisms can interface with language and with other specific domains (Nairne 1988, 1990). These mechanisms undergo development, which is likely to affect children's interpretations (for reviews, see Feigenson 2007; Cowan 2001; Courage and Cowan 2008). Importantly, this development does not require evidence from the linguistic input; essentially, a domain-general maturation account. At the same time, other processes may be more sensitive to specific input frequencies, as discussed above for potential cues for adjunct control in the input (for further discussion, see Van Dyke and Johns 2012; Omaki and Lidz 2015; Gerard 2016).

In both cases, children's non-adultlike interpretations are due to problems with deploying adultlike syntactic knowledge: first, antecedent retrieval may be more susceptible to interference, due to immature memory mechanisms; additionally, children may not have a strong enough link between the overt forms of tense or complementizers and the corresponding structures.[3]

These issues with deployment are not mutually exclusive. Also, their predictions can be tested in an experimental context (discussed further below). Regardless, they do not affect the arguments above about the lack of evidence in the input for attachment height or a c-commanding controller; for example, children are still likely to access non-adultlike interpretations of adjunct control in the input, regardless of the source of these non-adultlike interpretations. However, a deployment account aligns with the evidence in the input, in addition to explaining children's behavior.

4.3 *Predictions of a UG account of acquisition*

With innate properties as a part of UG, predictions are made for the linguistic input, as well as an experimental context.

[3] This second option is similar to the account proposed by Cairns et al (1994) in that adultlike behavior is achieved by forming adultlike mappings between lexical forms and abstract structure.

4.3.1 Predictions for the input

Although most types of adjunct control exhibit subject control, exceptions exist depending on various aspects of the dependency. To account for this variation, evidence must be available in the input in some form. For example, in (15), the controller is the main clause patient, rather than the subject:

(15) a. John$_1$ thanked Mary$_2$ for PRO$_{*1/2}$ running to the store.
b. John$_1$ was thanked by Mary$_2$ for PRO$_{1/*2}$ running to the store.

This exception with the complementizer *for* is observed across languages with the corresponding complementizer. This means that some aspect of the meaning of *for* is associated with control by the patient, or that evidence in the input is available for this exception.

To test this prediction, an additional corpus search was conducted for non-finite adjuncts with the complementizer *for*, using the same methods as described above. These counts are presented in Figure 6.

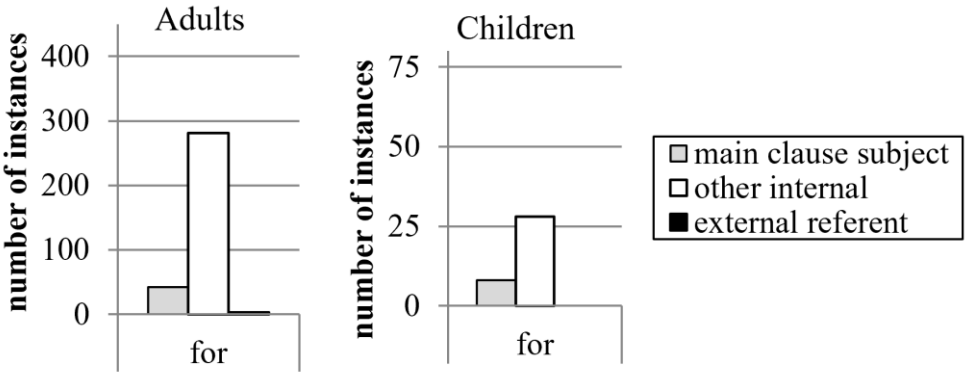


Figure 6. Frequency of antecedents in non-finite *for* by adults (speech to children) and children (speech by children) in CHILDES (note the different y axis scales).

The data here raise two main points. First, compared to the other non-finite complementizers, the adjuncts with *for* occur at a high frequency

(comparable to *without* and *instead of*), and should therefore be more salient than the lower frequency adjuncts.

Next, unlike the other non-finite complementizers, which occurred in the input with only subject control interpretations, an overwhelming majority of adjuncts with *for* have an object or other internal NP as the controller, as in the following examples:[4]

(16) a. Can you$_1$ scold Jennifer$_2$ for PRO$_{*1/2}$ stepping on the truck?

b. What did Aunt Carey$_1$ buy you$_2$ at the store for PRO$_{*1/2}$ being a good sharer?

c. You$_1$ yelled at him$_2$ today for PRO$_{*1/2}$ chewing your slippers.

d. I$_1$ have a little present for you$_2$ for PRO$_{*1/2}$ coming today.

If children are sensitive to different distributions of antecedents, this is the kind of striking contrast that might be relevant for acquisition. This would be in comparison to a contrast between strict subject control and e.g. a discourse bias for the subject interpretation, which would only be detectable in a minority of instances.

However, while some variety is observed within the instances of *for* adjuncts, 70% of the instances occurred in the frame 'thank you for ___ing,' as in:

(17) a. Thank you for helping me.

b. Thank you for letting Mommy finish her breakfast.

c. Thank you for carrying socks.

---

[4] The search of *for* followed by the string "ing" also returned utterances such as the following:

(i) They're not for eating.

(ii) Where're the songs for dancing?

(iii) This one's for something else.

(iv) Mommies are not for hitting.

These instances were not included in the analysis.

This frequent frame may start out as a larger chunk, to be linked later to the *for* non-finite frame. Meanwhile, the discourse contexts for the utterances in (16) strongly support a patient interpretation for the adjunct subject. These utterances, along with the instances with the patient as the subject, may provide the relevant evidence against strict subject control for *for* adjuncts.

This predicts, however, that similar evidence will be available in the input for other languages. It also predicts that children would treat *for* adjuncts like the other non-finite adjuncts until the relevant evidence is available. Alternatively, the meaning of *for* as a complementizer may be associated already with the patient antecedent, so that identifying the complementizer form-meaning mapping would be sufficient for acquisition; this would involve additional language-specific information to be specified in UG.

4.3.2 Predictions for the lab

Under a UG acquisition account, one key feature to acquire from the input is complementizer form, along with which complementizers select for non-finite adjuncts as opposed to finite adjuncts. This raises the question of whether a learner would have a default assumption to start with - similar to an initial parameter setting - to select for e.g. both finite and non-finite frames (like *after, before,* and *while*) or for non-finite frames only (like *without* and *instead of*). If there is such a default, then predictions may be made about children's behavior in an experimental context for low frequency complementizers, based on the corpus data discussed above.

First, if children have a default assumption that complementizers select for both finite and non-finite adjuncts, then their expectations about the likelihood of either frame should match the frequencies of higher frequency complementizers. For example, the adjuncts with *after, before,* and *while* as complementizers occurred much more frequently in the input in the finite frame than in the non-finite frame. Therefore, a default assumption for both frames should include a preference for finite adjuncts.

This prediction may be tested in an experimental context; in particular, for low frequency complementizers, children should be more likely to use a finite frame, even when this results in an ungrammatical

utterance. For example, the complementizer *despite* was also included in the corpus search discussed above, as in (18):

(18) John$_1$ called Mary$_2$ despite PRO$_{1/*2}$ going to the store.

Like *without* and *instead of*, *despite* can only appear in a non-finite frame; however, no utterances of adjunct control with *despite* were observed in CHILDES (for both North American and UK English corpora). Therefore, if children have a preference for finite clauses, then they should produce ungrammatical frames as in (19):

(19) *John$_1$ called Mary$_2$ despite he went to the store.

However, with no instances of *despite* (finite or non-finite) in the input, this raises the question of how a learner would ever receive evidence against finite frames.

Another option is for a default assumption that complementizers appear only in a non-finite frame. This avoids the issue of evidence for *despite*, predicting that children will produce a non-finite, grammatical version as in (18) in an experimental context. Of course, this alternative shifts the issue over to finite complementizers: for any low frequency finite complementizers, ungrammatical non-finite frames are predicted in an experimental setting, as in:

(20) *John$_1$ called Mary$_2$ because PRO going to the store.

In the corpus data, finite adjuncts were much more frequent than non-finite adjuncts. If there is a default setting for non-finite adjuncts, this contrast in frequency makes it particularly salient whether a given complementizer may select for finite adjuncts. Meanwhile, these selectional properties vary across languages; while for low frequency complementizers, the same behavior in an experimental context is predicted across languages, different distributions are predicted in the input to reflect this variation.

5. Discussion

This paper has considered the options for acquiring adjunct control. Although adjunct control is available in the input, this is not sufficient for acquiring the main syntactic properties of adjunct control. Observing instances of adjunct control directly may provide information about overt features in the dependency, but not abstract features like the correct attachment height of the adjunct or the controller as the closest c-commanding NP. Similar issues arise when considering the possibility of generalizing from other structures, which involve arbitrary assumptions about generalization.

Without evidence in the input for these key components of adjunct control in the input, they must be innate - considered here as principles in UG. This argument from the poverty of the stimulus instead involves a different type of evidence in the input for acquiring adjunct control, and makes further predictions about the input. The following sections consider the implications of this account - for control, for other dependencies, and for acquisition.

5.1 *Other types of control*

Accounting for the adjunct control as a dependency requires a syntactically defined locality constraint. This is supported by crosslinguistic judgments, as well as in experiments which control for the discourse context (Parker, Lago, and Phillips 2015; Broihier and Wexler 1995; Adler 2006; Gerard et al. 2018; but see Green 2018b). These instances of adjunct control are generally considered to be obligatory control in that they require a local antecedent. Meanwhile, non-obligatory control is also observed in temporal adjuncts, depending on a number of factors (Williams 1992; Landau 2015, 2017; Green 2018a). For example, a logophoric interpretation is available in the following sentences:

(21) a. Potatoes are tastier [after PRO boiling them].
b. The stairs were washed [before PRO entering the basement].
c. Those lemons tasted especially sour [after PRO eating sugar].

Therefore, if the abstract syntactic components of obligatory control are innate, more details are needed to account for non-obligatory control.

Previous research on children's acquisition has generally focused on obligatory control. In CHILDES, a few instances of logophoric control are observed:

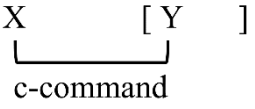
(22) a. The pastry has got to rest for a little while in the fridge before PRO rolling it out.
b. Mine [=tricycle] will go sailing the seven seas without PRO pushing a button.

The utterance in (22a) is produced by a mother, while the utterance in (22b) is produced by a child (aged 4;11). However, it is also possible that these instances would not be differentiated from those of the type $data_2$ from Figure 1, which would be interpreted as obligatory control but with a non-adultlike interpretation. Future research will further examine these implications for acquiring obligatory and non-obligatory control.

5.2 *Other dependencies*

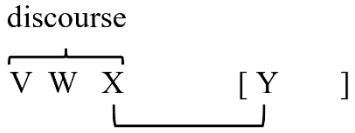
This paper discusses the acquisition of adjunct control based on a hierarchical relation (c-command by the controller) and attachment height. In addition to adjunct control, other dependencies are also defined in terms of hierarchical relations, so much of the logic discussed here may be applied more generally.

For example, for any referential dependency, an antecedent must be identified to resolve the dependency. Consider a syntactic dependency between X and Y, where the grammatical antecedent may be identified by some constraint (e.g. c-command and/or locality):

(23)


If the relevant constraint has not yet been acquired, then an alternative strategy is needed to resolve the dependency; for example, by retrieving an antecedent from the discourse:

(24)



Additionally, there must be evidence available in the input to (eventually) acquire the relevant syntactic constraint. Otherwise, without this evidence, some aspect of the dependency must be available in UG; this will make further predictions similar to adjunct control about factors like exceptions, experimental contexts, etc.

Languages vary in their inventories of syntactic dependencies, with some dependencies observed more universally than others. Positing a domain-specific feature in UG may account for more widely observed dependencies, while evidence is needed in the input in other cases. Arguments identifying which features are in UG often (reasonably) appeal to this universality, or lack thereof; this paper is concerned also with the transparency of a given feature in the input: for abstract properties which are not directly observable from the linear input, evidence for these properties may be more elusive, even when the relevant structures are available in the input. Attachment height and c-command are examples of such properties (with the same logic for locality in other frameworks).

Importantly, the contrast between these abstract properties and the language-specific input forms represents a continuum, rather than a binary distinction. While control is generally observed across languages, other types of dependencies are observed widely but not universally, such as long-distance anaphors or logophors (Koster and Reuland 1991; Hyams and Sigurjonsdottir 1990; Chien, Wexler, and Chang 1993). These involve abstract properties, but also require evidence in the input about the relevance of these properties. Additionally, evidence is needed for exceptions to a given dependency. These are observed across languages, like with *for* as a complementizer as discussed above. Exceptions are also observed within languages, as in (25):

(25) a. O João$_1$ cumprimentou a Maria$_2$ depois de
the João$_1$ greeted the Maria$_2$ after of
PRO$_{1/*2}$ entrar na sala
PRO$_{1/*2}$ enter in-the room.
"João greeted Maria after entering the room."

b. Que Mulher$_2$ é que o João$_1$ cumprimentou t$_2$
which woman is that the João$_1$ greeted
depois de PRO$_{\mathbf{1/2}}$ entrar na sala
after of PRO$_{1/2}$ enter in-the room.
"Which woman did João greet after entering the room?"
adapted from Nunes (2014)

In Brazilian and European Portuguese, obligatory subject control is observed in non-finite clauses like in (25a). However, both subject control and object control are possible in contexts when the matrix object undergoes wh-movement, as in (25b).

This difference between English and Portuguese is accounted for formally in terms of the features on the wh-phrase in English compared to Portuguese (Nunes 2014). Meanwhile, children acquiring Portuguese must receive evidence from the input for this contrast. Of the instances of adjunct control in CHILDES in English, 12 (or 1.4%) included wh-movement. The moved wh-word was not a plausible antecedent of PRO in any of these instances. This is consistent with the judgments in English, as evidence should not be available for object control with wh-movement in English. In contrast, this evidence should be available in the input for Portuguese, either directly (based on wh-questions in the input) or indirectly, via another form of evidence. Depending on what this evidence is, there may be further implications for the content of UG.

Thus, this logic for identifying evidence in the input - used in this paper for adjunct control - may also be applied for cross-linguistic variation at different scales. It can help to identify the evidence available for abstract properties, and the variation in these properties.

5.3 *Role of the argument of the poverty of the stimulus*

This paper presents an argument from the poverty of the stimulus that the abstract components of adjunct control are innate. Evidence for these components does not occur in the input, so they must be available from another source. If attachment height and the controller are part of UG, then common features of control across languages may be explained without requiring redundancy in the input. This paper also discussed further predictions about the evidence in the input for acquiring control, and for cross-linguistic variation.

More broadly, based on the type of evidence that is *not* available and because these features of control are *not* learned, the conclusions about evidence in the input are applicable to linguistic dependencies more generally: if the actual elements of a dependency are not reliable for inferring the properties of the dependency, then a different form of evidence is needed for these properties. This was the case for adjunct control, as children's non-adultlike interpretations of adjunct PRO were likely to provide incorrect evidence about the adult grammar. Similarly, non-adultlike interpretations have also been observed for other types of anaphora (Chien and Wexler 1990; McKee 1992; for a review see Conroy et al. 2009), as well as A movement (Manzini and Wexler 1987; Orfitelli 2012; Mateu 2016, inter alia) and A-bar movement (Tavakolian 1981; Friedmann, Belletti, and Rizzi 2009; Adani et al. 2010, inter alia; but see Hamburger and Crain 1982; Gagliardi, Mease, and Lidz 2016).

For many of these general phenomena, innate components have been proposed, based on the poverty of the stimulus. Meanwhile, children's non-adultlike behavior is often accounted for by a non-adult grammar. These accounts may achieve descriptive adequacy for children's non-adultlike behavior; however, if evidence is not available in the input for the non-adult grammar *and* for the transition to the adult grammar, then this casts doubt on the explanatory adequacy of the grammar. If both forms of evidence are not available, then either a different non-adult grammar or extragrammatical sources are needed to account for children's behavior.

5.4 *Conclusion*

This paper considered how adjunct control is acquired and compared different sources of evidence in the linguistic input. These options did not provide evidence for the key grammatical components of adjunct control, suggesting that these components are innate, with other more overt forms of evidence in the input. Future research will further investigate the predictions of this evidence, as well as the more general implications for the content of UG.

References

Adani, Flavia, Heather KJ Van der Lely, Matteo Forgiarini, and Maria Teresa Guasti. 2010. Grammatical Feature Dissimilarities Make Relative Clauses Easier: A Comprehension Study with Italian Children. *Lingua* 120(9):2148–2166.

Adler, Allison Nicole. 2006. Syntax and Discourse in the Acquisition of Adjunct Control. Massachusetts Institute of Technology.

Akhtar, Nameera, Maureen Callanan, Geoffrey K. Pullum, and Barbara C. Scholz. 2004. Learning Antecedents for Anaphoric One. *Cognition* 93(2):141–45.

Arnold, Jennifer E. 2001. The Effect of Thematic Roles on Pronoun Use and Frequency of Reference Continuation. *University of Pennsylvania Working Papers in Linguistics* 6(3):209–35.

Baker, C. L. 1979. Syntactic Theory and the Projection Problem. *Linguistic Inquiry* 10(4):533–81.

Berwick, Robert C. 1985. The Acquisition of Syntactic Knowledge. MIT Press.

Broihier, Kevin and 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, and Michelle Rapp. 1994. A Longitudinal Study of Principles of Control and Pronominal Reference in Child English. *Language* 70(2):260–288.

Chien, Yu-Chin and Kenneth Wexler. 1990. Children's Knowledge of Locality Conditions in Binding as Evidence for the Modularity of Syntax and Pragmatics. *Language Acquisition* 1(3):225–295.

Chien, Yu-Chin, Kenneth Wexler, and Hsing-Wu Chang. 1993. Children's Development of Long-Distance Binding in Chinese. *Journal of East Asian Linguistics* 2(3):229–59.

Chomsky, Noam. 1965. *Aspects of the Theory of Syntax*. Vol. 11. MIT press.

Chomsky, Noam. 1981. Lectures on Government and Binding: The Pisa Lectures. Walter de Gruyter.

Conroy, Anastasia, Eri Takahashi, Jeffrey Lidz, and Colin Phillips. 2009. Equal Treatment for All Antecedents: How Children Succeed with Principle B. *Linguistic Inquiry* 40(3):446–486.

Courage, Mary and Nelson Cowan. 2008. *The Development of Memory in Infancy and Childhood*. Psychology Press.

Cowan, Nelson. 2001. The Magical Number 4 in Short-Term Memory: A Reconsideration of Mental Storage Capacity. *Behavioral and Brain Sciences* 24(01):87–114.

Crain, Stephen and Cecile McKee. 1985. The Acquisition of Structural Restrictions on Anaphora. Pp. 94–110 in *Proceedings of NELS*. Vol. 15.

Crain, Stephen and Rosalind Thornton. 1998. Investigations in Universal Grammar: A Guide to Experiments on the Acquisition of Syntax and Semantics. MIT Press.

Feigenson, Lisa. 2007. Continuity of Format and Computation in Short-Term Memory Development. *Short-and Long-Term Memory in Infancy: Taking the First Steps towards Memory* 51–74.

Friedmann, Naama, Adriana Belletti, and Luigi Rizzi. 2009. Relativized Relatives: Types of Intervention in the Acquisition of A-Bar Dependencies. *Lingua* 119(1):67–88.

Gagliardi, Annie, Tara M. Mease, and Jeffrey Lidz. 2016. Discontinuous Development in the Acquisition of Filler-Gap Dependencies: Evidence from 15- and 20-Month-Olds. *Language Acquisition* 23(3):234–60.

Gerard, Juliana. submitted. The Extragrammaticality of the Acquisition of Adjunct Control. 35.

Gerard, Juliana. 2016. The Acquisition of Adjunct Control: Grammar and Processing.

Gerard, Juliana, Jeffrey Lidz, Shalom Zuckerman, and Manuela Pinto. 2017. Similarity-Based Interference and the Acquisition of Adjunct Control. *Frontiers in Psychology* 8.

Gerard, Juliana, Jeffrey Lidz, Shalom Zuckerman, and Manuela Pinto. 2018. The Acquisition of Adjunct Control Is Colored by the Task. *Glossa: A Journal of General Linguistics* 3(1).

Gold, E. Mark. 1967. Language Identification in the Limit. *Information and Control* 10(5):447–474.

Goodluck, Helen. 1981. Children's Grammar of Complement-Subject Interpretation. Pp. 139–66 in *Language acquisition and linguistic theory*, edited by S. Tavakolian. Cambridge, MA: The MIT Press.

Goodluck, Helen and Dawn Behne. 1992. Development in Control and Extraction. Pp. 151–71 in *Theoretical issues in language acquisition*, edited by J. Weissenborn, H. Goodluck, and T. Roeper. Hillsdale, NJ: Lawrense Erlbaum.

Green, Jeffrey. 2018a. Adjunct Control: Syntax and Processing. PhD Thesis, University of Maryland.

Green, Jeffrey. 2018b. Non-Subject Control of Temporal Adjuncts. in *Proceedings of CLS 54*.

Grimshaw, Jane and Steven Pinker. 1989. Positive and Negative Evidence in Language Acquistion. *Behavioral and Brain Sciences* 12(2):341–342.

Guasti, Maria Teresa. 2017. Language Acquisition: The Growth of Grammar. MIT Press.

Haegeman, Liliane. 2000. Adult Null Subjects in Non Pro-Drop Languages. *The Acquisition of Syntax: Studies in Comparative Developmental Linguistics* 129–169.

Hamburger, Henry and Stephen Crain. 1982. Relative Acquisition. *Language Development* 1:245–274.

Harrigan, Kaitlyn, Valentine Hacquard, and Jeffrey Lidz. 2019. Hope for Syntactic Bootstrapping. *Language* 95(4):642–82.

Hart, Betty and Todd R. Risley. 1995. Meaningful Differences in the Everyday Experience of Young American Children. Paul H Brookes Publishing.

Heinz, Jeffrey and Jason Riggle. 2011. Learnability. *The Blackwell Companion to Phonology, 5 Volume Set* 2:54.

Holmberg, Anders, Aarti Nayudu, and Michelle Sheehan. 2009. Three Partial Null-Subject Languages: A Comparison of Brazilian Portuguese, Finnish and Marathi*. *Studia Linguistica* 63(1):59–97.

Hsu, Jennifer Ryan, Helen Smith Cairns, and Robert W. Fiengo. 1985. The Development of Grammars Underlying Children's Interpretation of Complex Sentences. *Cognition* 20(1):25–48.

Huang, C. T. James. 1984. On the Distribution and Reference of Empty Pronouns. *Linguistic Inquiry* 531–574.

Hyams, Nina and Sigridur Sigurjonsdottir. 1990. The Development of Long-Distance Anaphora: A Cross-Linguistic Comparison With Special References to Icelandic. *Language Acquisition* 1(1):57–93.

Kehler, Andrew and Hannah Rohde. 2013. A Probabilistic Reconciliation of Coherence-Driven and Centering-Driven Theories of Pronoun Interpretation. *Theoretical Linguistics* 39(1–2):1–37.

Kempe, Vera and Brian MacWhinney. 1999. Processing of Morphological and Semantic Cues in Russian and German. *Language and Cognitive Processes* 14(2):129–71.

Koster, Jan and Eric Reuland. 1991. *Long Distance Anaphora*. Cambridge University Press.

Landau, Idan. 2015. *A Two-Tiered Theory of Control*. MIT Press.

Landau, Idan. 2017. Adjunct Control Depends on Voice. *A Pesky Set: Papers for David Pesetsky* 93–102.

Lidz, Jeffrey and Annie Gagliardi. 2015. How Nature Meets Nurture: Universal Grammar and Statistical Learning. *Annu. Rev. Linguist.* 1(1):333–353.

Ler, Gereon M. 1995. *A-Bar Syntax: A Study in Movement Types*. Walter de Gruyter.

Lust, Barbara, Julie Eisele, and Reiko Mazuka. 1992. The Binding Theory Module: Evidence from First Language Acquisition for Principle C. *Language* 68(2):333–58.

MacWhinney, Brian. 2000. *The CHILDES Project: The Database*. Vol. 2. Psychology Press.

Manzini, M. Rita and Kenneth Wexler. 1987. Parameters, Binding Theory, and Learnability. *Linguistic Inquiry* 18(3):413–44.

Mateu, Victoria Eugenia. 2016. Intervention Effects in the Acquisition of Raising and Control: Evidence from English and Spanish. PhD Thesis, UCLA.

McDaniel, Dana, Helen Smith Cairns, and Jennifer Ryan Hsu. 1991. Control Principles in the Grammars of Young Children. *Language Acquisition* 1(4):297–335.

McKee, Cecile. 1992. A Comparison of Pronouns and Anaphors in Italian and English Acquisition. *Language Acquisition* 2(1):21–54.

Nairne, James S. 1988. A Framework for Interpreting Recency Effects in Immediate Serial Recall. *Memory & Cognition* 16(4):343–352.

Nairne, James S. 1990. A Feature Model of Immediate Memory. *Memory & Cognition* 18(3):251–269.

Newport, Elissa L. 2016. Statistical Language Learning: Computational, Maturational, and Linguistic Constraints. *Language and Cognition* 8(3):447–61.

Nunes, Jairo. 2014. Moving to the Edge: The Case of Adjunct Control in Portuguese.

Omaki, Akira and Jeffrey Lidz. 2015. Linking Parser Development to Acquisition of Syntactic Knowledge. *Language Acquisition* 22(2):158–192.

Orfitelli, Robyn Marie. 2012. Argument Intervention in the Acquisition of A-Movement. PhD Thesis, UCLA.

Parker, Dan, Sol Lago, and Colin Phillips. 2015. Interference in the Processing of Adjunct Control. *Frontiers in Psychology* 6.

Pinker, Steven. 1979. Formal Models of Language Learning. *Cognition* 7(3):217–283.

Pinker, Steven. 2009. Language Learnability and Language Development, with New Commentary by the Author. Vol. 7. Harvard University Press.

Pinker, Steven. 2013. Learnability and Cognition: The Acquisition of Argument Structure. MIT press.

Sundaresan, Sandhya. 2014. Making Sense of Silence: Finiteness and the (OC) pro vs. pro Distinction: Commentary on Kissock (2013). *Natural Language & Linguistic Theory* 32(1):59–85.

Sutton, Megan. 2015. Competence and Performance in the Development of Principle C.

Tavakolian, Susan. 1981. The Conjoined-Clause Analysis of Relative Clauses. *Language Acquisition and Linguistic Theory* 167:187.

Tenenbaum, Joshua B. 1999. Bayesian Modeling of Human Concept Learning. Pp. 59–68 in *Advances in Neural Information Processing*

Systems 11, edited by M. J. Kearns, S. A. Solla, and D. A. Cohn. MIT Press.

Tenenbaum, Joshua B. and Thomas L. Griffiths. 2001. Generalization, Similarity, and Bayesian Inference. *Behavioral and Brain Sciences* 24(04):629–640.

Van Dyke, Julie A. and Clinton L. Johns. 2012. Memory Interference as a Determinant of Language Comprehension: Interference in Comprehension. *Language and Linguistics Compass* 6(4):193–211.

Wexler, Kenneth and Yu-Chin Chien. 1985. The Development of Lexical Anaphors and Pronouns. *Papers and Reports on Child Language Development* 24:138–49.

Williams, Edwin. 1992. Adjunct Control. Pp. 297–322 in *Control and Grammar*, *Studies in Linguistics and Philosophy*, edited by R. K. Larson, S. Iatridou, U. Lahiri, and J. Higginbotham. Springer Netherlands.

Wu, Guobin. 1992. Discourse Anaphora in Chinese: A Rhetorical Predicate Account.

Xu, Fei and Joshua B. Tenenbaum. 2007. Word Learning as Bayesian Inference. *Psychological Review* 114(2):245–72.