The Acquisition of Syntax

Nina Hyams   Robyn Orfitelli
UCLA       University of Sheffield

Introduction

How children acquire their native language(s) is a central concern of linguistic theory and cognitive science more generally. There is a clear consensus that humans are specifically adapted to the task of language acquisition and that language development is driven by innate capacities and shaped by the environment. Though there are many different views on the nature of the innate mechanisms, as well as the degree of environmental influence, any explanation must reckon with the fact that the language/grammar that children acquire is vastly underdetermined by the linguistic input they receive. Children are exposed to adult utterances, not abstract grammatical structures, yet acquire a hierarchically organized system of categories and rules. Children hear only a finite number of adult utterances, yet have the capacity to produce and understand sentences way beyond this restricted set – including many that are not grammatical in the adult “input” language. Lastly, children are provided with “positive” exemplars (possible sentences in his language), yet somehow come to also know what is not possible in their language. In short, every typically developing child develops a rich and highly abstract system of rules that constitutes the grammar of their language. They do this in a very short period of time, based on rather limited evidence, without the benefit of explicit correction, instruction, or information about ungrammaticality, and much of what they come to know is not transparently exemplified in the language of the adults around them. The acquisition problem couched in these terms is often referred to as the logical problem of language acquisition (LPLA) (Chomsky, 1965; Baker & McCarthy, 1981).

The LPLA can be illustrated using the following sentences:

(1) a. Ernie looks like he rode his bike to school today.
   b. It looks like Ernie rode his bike to school today.

The sentences in (1) seem to be simple paraphrases of each other, and indeed, in certain situations either sentence could be felicitously uttered. For example, imagine Ernie is standing next to his bicycle in front of a school building, wearing a bicycle helmet. In that case we might equally well utter sentence (1a) or (1b). However, now imagine we’re looking at a classroom. In it we see Ernie’s desk and his bicycle helmet is under the desk, but the classroom is empty because all the children are at recess. In this case sentence (1b) is still a perfectly appropriate utterance, but (1a) is decidedly odd, because we have no visible evidence of Ernie himself. The contrast is subtle. Both sentences are fully grammatical, but they are not equivalent. They have different evidential requirements. Clearly, this is not something children are instructed on.
Moreover, it is unlikely that they have any kind of (negative) evidence which would lead them to know that (1a) is not felicitous in certain contexts. Yet, children as young as two-years old use both kinds of sentences in the appropriate situations (Rett & Hyams, 2013; Rett, Hyams, & Winans, 2012).

There are many such cases in natural language, sentences that should be possible but are not because they are blocked by some grammatical constraint. In the examples in (1) the constraint is semantic in nature – the determiner phrase (DP) subject of the matrix clause must be the perceptual source for the assertion (Asudeh & Toivonen, 2012). In other cases, such as those discussed below, the constraint is syntactic. In all cases, the linguistic evidence for the constraint that is available to the child is slim or non-existent.

Most linguists assume that the solution to the LPLA lies in the theory of Universal Grammar (UG), a set of innately specified grammatical principles that provides a blueprint for human languages, and that “guides” the child’s language development by restricting his grammar-forming options. While the exact form of UG is open to empirical investigation, it will ideally explain those properties that are invariant across languages, as well as cross-linguistic differences that are easily describable within a restricted parameter space. Most crucially, UG must be sufficiently articulated so that - together with the available input evidence – it provides an acquisition path to any (and all) target grammars.

But the child’s path to the target grammar is not without some curves and potholes, and it is not instantaneous. Children go through fairly well-defined stages, including the well-studied null subject and root infinitive stages discussed later in this chapter, in which their productions deviate from the adult’s in systematic ways. Similarly, they do not necessarily have an adult understanding of all sentence structures, among these certain types of passives, raising and control structures. We will discuss each of these phenomena below. A comprehensive picture of the child’s grammatical development means understanding the initial state -- the principles that constrain development -- and also the nature of his linguistic “errors” and stages. We can refer this as the developmental problem of language acquisition. An important aspect of the developmental problem is determining which properties of early language are due to representational differences between the child and the adult and which properties result from the child’s more limited language processing resources. In the latter case the child may have the adult grammatical representation but be unable to produce or understand a construction due to limitations of working memory, sentence planning, articulatory control, and so on. The interacting effects of linguistic competence and performance (Chomsky, 1965) are nowhere better illustrated than in child language.

In the following sections we briefly describe several aspects of children’s syntactic development. Given the limitations of space we do not intend this to be a comprehensive review. Rather, we have chosen facets of grammar acquisition that have been especially well studied and which represent different developmental stages and processes. Additionally, we focus on areas that allow us to illustrate the different kinds of data that inform theories of children’s grammars – naturalistic production data as well as experimental results. Finally, in line with the considerations just discussed, we chose topics that illustrate the challenges in deciding between competence-based or performance-based accounts of particular developmental phenomena.
Early multi-word utterances: The root infinitive stage

At around two years of age children begin to produce multi-word utterances. This might properly be considered the first stage of syntax acquisition in production, although comprehension of syntax begins earlier (see for example Golinkoff & Hirsch-Pasek, 1996). Children’s earliest utterances are short (typically 2-3 words) and consist largely of open class lexical elements (e.g. nouns, verbs, adjectives), with relatively few closed class functional elements (e.g. articles, auxiliaries, prepositions). This early language is traditionally referred to as ‘telegraphic speech’ (Brown, 1973; Brown & Fraser, 1964). Morphological development in the telegraphic stage is described in detail by Deen in Chapter X of this volume; here we will limit our discussion to one of its most notable morphosyntactic features: root infinitives (RIs; Rizzi, 1993/4), non-adultlike sentences in the main verb occurs in its infinitival form. (Because children produce RIs alongside finite forms, the stage is also referred to as the optional infinitive stage (Wexler, 1994)). Rates of RI production vary by child and language, ranging from 26% to 61% of all (finite and non-finite) verbal utterances (Hoekstra & Hyams, 1998).

In (2) we provide examples of RIs and finite verbs from several of the languages that show an RI stage. We focus on languages other than English because, strictly speaking, English has no morphological infinitive. What passes for an infinitive, a bare verb, as in (2f), can also result from the simple dropping of the finite ending (See Song, Sundara, & Demuth, 2009 for evidence that phonological factors are at play in the production of English bare verbs).

\[(2)\]
\[\begin{align*}
\text{a.} & \quad \text{Hun sove.} \quad \text{(Danish: Jens 2;0, Hamann & Plunkett, 1998)} \\
& \quad \text{She sleep-INF} \\
& \quad \text{‘She sleeps.’} \\
\text{b.} & \quad \text{Dormir petit bébé.} \quad \text{(French: Daniel 1;11, Pierce, 1989)} \\
& \quad \text{sleep-INF little baby} \\
& \quad \text{‘Little baby sleep.’} \\
\text{c.} & \quad \text{Earst kleine boekje lezen.} \quad \text{(Dutch: Hein 2;6, Haegeman, 1995)} \\
& \quad \text{First little book read-INF} \\
& \quad \text{‘First (I/we) read little book.’} \\
\text{d.} & \quad \text{S[ch]okolade holen.} \quad \text{(German: Andreas 2;1; Kramer, 1993)} \\
& \quad \text{chocolate get-INF} \\
& \quad \text{‘I got chocolate (?)’} \\
\text{e.} & \quad \text{Ty mama pomogat’} \quad \text{(Russian: Vavara 2;0,1 Brun & Babyonyshev, 2006)} \\
& \quad \text{You.nom mama.dat help-INF}
\end{align*}\]
‘You (=I) help mommy’

f. Cromer wear glasses  (English: Eve 2;0, Brown, 1973)

Additionally, the co-occurrence of non-finite and finite forms during the same period strongly suggests that RIs are not simply a “default” form that children use when they do not know inflected forms (cf. Blom, 2003). Indeed, there are many properties that distinguish the two clause types. The various ways in which RIs differ from finite clauses are outlined in (3):

(3) RIs typically:
   (i) have infinitival morphology in languages that with an infinitival form (e.g. 2a-e)
   (ii) do not occur with subject clitics in French.
   (iii) do not occur as copula or auxiliaries; the copula and auxiliaries are always finite.
   (iv) appear in sentence final position in OV/V2 languages such as Dutch and German (cf. 2c-d) whereas finite verbs occur in second position.
   (v) do not occur in wh-questions; over 97% of wh-questions are finite in Dutch, German, Swedish, French.
   (vi) do not occur with unambiguous (i.e. object, adverb, prepositional phrase) topics in languages with topicalization such as Dutch, German, Icelandic.
   (vii) occur only with eventive verbs, not with statives.
   (viii) occur overwhelmingly with null subjects (NS) (e.g. 2c-d); the NS rate in RIs is 75% to 90% compared to 8 to 32% NS rate in finite clauses.
   (ix) do not occur (or occur at much lower rates) in null subject languages such as Italian and Spanish; in these languages fewer than 15% of verbal utterances are RIs and typically the rate is in the single digits.

The properties associated with RIs are reviewed in detail by Hoekstra and Hyams (1998). English bare verbs (cf. 2f) fail to show some of the properties in (3) that would be expected if they are truly non-finite. For example, bare verbs occur in wh-questions (3v), with stative and eventive verbs (3vii), and the rate of null subjects is roughly the same for bare and finite verbs (3viii).

Various grammatical explanations for the RI stage have been proposed. Of particular interest is Rizzi’s (1994) hypothesis that the grammar of young children need not project a full complementizer phrase (CP) structure, but instead can generate structures in which any maximal projection (XP) constitutes a licit root. RIs result when the structure is truncated below the tense phrase (TP), arguably at the verb phrase (VP); finite clauses result from a full CP projection, or given assumptions of an expanded left periphery (e.g. Rizzi 1997), whatever projection is taken to be the highest. The truncation hypothesis, schematized below, successfully accounts for many, if not all of the properties in (3).
Without an inflectional system (including tense, agreement, and auxiliaries) there is no finite morphology, subject clitics (as in French), or auxiliary verbs; without a CP layer there can be no *wh*-questions, topicalizations or movement of the verb to C (as in German and Dutch). Although Rizzi does not directly address the eventivity effect in (3vii), it is easily accountable under a theory in which eventive and stative verbs have a distinct syntax, with the latter involving higher functional structure, as proposed by Borer (2005) and elsewhere under different assumptions. The NS facts in (3viii) and a lack (or greatly reduced) RI stage in null subject languages such as Italian and Spanish (3ix) require some additional assumptions, but can also be handled by Rizzi’s hypothesis.

Rizzi assumes that truncation is a UG-compliant option that also occurs in particular registers of adult grammars (4). For adults, special pragmatic or semantic factors are often required, but it would seem that children can avail themselves of the truncation option without such pragmatic licensing.

   b. Hier geen fietsen plaatsen! (Dutch)
      here no bicycles place-inf
      ‘Don’t put bicycles here’

The majority of accounts of the RI stage are competence-based; that is to say, they propose that the child’s grammar differs from the adult grammar in licensing RIs. Alternatives to the truncation theory include, for example, the Null Aux hypothesis (Boser et al 1992) and the Agreement-Tense Omission Model (ATOM, Schutze & Wexler 1996), both of which propose a difference between the children’s grammar and the target. In recent years, however, an increasing number of researchers have appealed to performance limitations as a possible explanation for the RI stage. According to Dye (2011), French-speaking children in the RI stage are attempting to produce a periphrastic structure (auxiliary/modal + non-finite verb) as in (5a), but often fall short of the target and produce a phonologically reduced form of the auxiliary, as in (5b) or delete it altogether, as in (5c), giving the appearance of a non-finite root clause.

(5) a. Je veux jouer avec le château fort moi
    I want.1st p. play.inf. with the fortress me
    ‘I want to play with the fortress.’
b. J’v’ouer avec le tast fort moi

c. Rentre na maison

return the home

(= il est rentrè/va rentrer à la maison – ‘he entered/is going to enter the house’)

Dye presents interesting arguments in support of the phonological reduction hypothesis including acoustic evidence of ‘covert’ auxiliaries in some RIs. What is left largely unexplained by this kind of account are several of the findings in (3). For example, why is there not a robust RI stage in Italian and Spanish, as these languages are closely related to French? Why do we fail to find RIs in topicalized sentences and wh-questions in V2 languages? Finally, why are RIs restricted to eventive predicates? Although a processing-based account might in theory be able to explain these patterns, further work remains to be done.

Rather than an exclusively grammatical or performance account, it is possible that RIs result from an interaction of grammatical and performance factors. For example, Rizzi’s grammatical truncation analysis is also consistent with a hypothesis in which children produce reduced (RI) structures under the pressure of more limited production or processing resources. In effect, they take advantage of a grammatical option that imposes a reduced computational load. Indeed, the notion that certain child language phenomena arise from a complex interaction of grammatical and processing factors has also been proposed by Rizzi (2005) as an explanation for another salient property of early language – the null subject stage. We turn to children’s null subjects in the next section.

**Null subjects in early language**

Children in the telegraphic stage also omit subjects, even where it is not a grammatical option in the adult language. This behavior not only occurs during a similar time period to the RI stage, but also in an overlapping set of languages (6).


   b. Ikke tøre traktor. (Danish: Jens 2;0, Hamann & Plunkett, 1998)

   not drive tractor

   c. Eerst kleine boekje lezen. (Dutch: Hein 2;6, Haegeman, 1995)

   first little book read

   d. Va sous la tabouret. (French: Philippe 2;2, Suppes et al 1973)

   goes under the stool

   e. Hubsauber putzn. (German: Andreas 2;1, Kramer, 1993)

   helicopter clean
Like RIs, subject omission is optional; during the same period in which children drop subjects in some sentences, they also produce them in many others (Hyams, 1986). The frequency of subject omission varies across individual children, across languages, and within the same child across time (Similar questions arise vis á vis the ‘null object’ stage, in which children acquiring Romance languages omit objects in environments where adults would use a clitic pronoun. See Mateu (2014) for discussion).

Explanations of the null subject (NS) stage fall into two broad categories. Competence-based accounts take children’s erroneous productions to reflect a difference between the child and adult grammars (Hyams, 1986, 1992; Hyams & Wexler, 1993; Rizzi, 1994, 2000; Yang, 2002, 2004, among others), although these accounts differ in various ways (for example, Rizzi’s account implicates truncation and Yang’s contains an important statistical component.). Conversely, performance accounts hold that subject omission is due to extra-syntactic factors (L. Bloom, 1970; P. Bloom, 1990; Gerken, 1991; Valian, 1991). On this view, children’s syntactic representations include an overt subject, just as in the adult representation, but the subject is dropped in production because of these other influences. Finally, as noted in the previous section, Rizzi (2005) has proposed a hybrid model in which both competence and performance factors conspire to explain children’s subject omission. Given the parametric option of null versus overt subjects – an option not available in the adult grammar – children will choose the computationally least expensive option.

While competence- and performance-based accounts of the NS stage each predict that children will drop subjects in production (for a discussion of other issues related to comprehension versus production, see Hyams 2011 and McKee, McDaniel, & Garrett, this volume), they categorically differ in their predictions of how children will comprehend NS sentences. Competence accounts predict that a child who produces NS sentences allows them to have a declarative interpretation. Performance accounts, however, predict that children should interpret NS utterances in an adult manner, which in English means only as imperatives.

Orfitelli and Hyams (2012) tested children’s comprehension of NS sentences with a modified version of the Truth-Value Judgment (TVJ) experiment (Crain & McKee, 1985; Crain & Fodor, 1993). The task tested whether they would permit an NS sentence to have a declarative interpretation in addition to (or instead of) an adult-like imperative one.

Participants saw pairs of pictures depicting two sets of children. The first picture showed two older children engaged in a particular activity, such as drawing a picture or playing with blocks, while the second picture always showed two younger children in close proximity to the relevant items (e.g., paper and crayons or blocks) but not interacting with them. Participants were told that the four children have the same babysitter, and that while the older children are old enough to choose their own activities without permission, the younger children had to wait for the babysitter to tell them to engage in the activity in question. This set up a mood-based dichotomy: it is appropriate to use a declarative sentence in reference to the picture of the older children, but not the younger children; conversely, it is appropriate to use an imperative sentence
in reference to the picture of the younger children, but not the older children. Strictly speaking, judging the appropriateness of imperative sentences does not involve computing a ‘truth-value,’ however, these items nonetheless test participants’ *comprehension* of the experimental scenarios, which sharply distinguishes this task from metalinguistic receptive tasks such as the grammaticality/acceptability judgment task.

After each story, one of the two pictures was removed, and an observing puppet made a comment about the remaining picture. Some of these comments took the form of NS sentences (7). The participant was asked to tell the puppet if his comment was correct or incorrect.

(7) Play with blocks.

If children accept NS sentences when paired with the ‘older-children’ scenario, it indicates that they permit a declarative syntactic structure for these sentences, which is fundamentally different from the target adult grammar. This is exactly what Orfitelli and Hyams found. Children in the NS stage accept both declarative and imperative interpretations for NS sentences, but stopped allowing the non-adult declarative interpretation at approximately 3½ years. This shift mirrors the one that has been seen in production, and indeed, we also found a near-perfect within-subject correspondence between NS production and non-adult comprehension.

This connection is not expected on a performance account, but is explicitly predicted by a grammatical account, as both behaviors would arise from the same root cause: the child’s different grammar. A performance account that attributes NSs solely to planning or production limitation (e.g. Bloom, 1991; Valian, 1991) cannot explain why children accept NSs in comprehension, and in particular why they do so under a non-adult (=declarative) interpretation. However, the pattern of results in not entirely straightforward. If children’s grammar licenses an ambiguity between imperative and declarative readings for NS sentences, we might expect them to answer *true* to all NS items, regardless of scenario. Instead, the children appeared to resolve the ambiguity in one direction or the other, and only then consider whether or not this mood matched the scenario that had just been presented. Approximately 50% of their answers seemed to indicate an adult, imperative interpretation for the NS items, while the other 50% of their answers seemed to indicate a non-adult, declarative interpretation.

The participants’ justification responses support this conjecture. On the imperative and declarative control conditions, children provided justifications that unambiguously indicated the correct mood; for example, justifications for imperative controls used phrases such as “supposed to” or “because he is telling them to do X,” but the justifications for the declarative control items were always simple declaratives. On the NS condition, participants’ responses matched these two patterns. When answering correctly, their justifications always resembled their responses to the imperative controls (8a), which is expected if the children were assigning imperative mood to the NS sentence. Conversely, when answering incorrectly, their justifications matched those for the declarative controls, as if they were assigning declarative mood to the sentence (8b).

(8) a. He needs to tell the younger kids to put on socks. –S., 3;3

Child correctly answers true to “Put on socks.”
Child incorrectly answers false to “Eat a cookie.”
b. Because they aren’t eating a cookie –E., 3;10

For any given NS test item, why were children allowing only a declarative or imperative interpretation, given the evidence that their grammar licenses both? This may be where processing factors may come into play. Young children are noted to have difficulty using context to disambiguate ambiguity in language comprehension tasks. Not only does this behavior show up in lexical ambiguity resolution (Swinney & Prather, 1989), but also in sentence-level attachment ambiguity (Trueswell et al 1999) and scopal ambiguity (Syrett & Lidz, 2005). The consequence of this is that in contrast to adults, who are able to revise incorrect parses to reflect additional contextual and other information, children appear to be bound by their original parse.

If NS sentences present an ambiguity for children, then when attempting to assign a meaning to a NS utterance, the children have to decide between a declarative and imperative representation for the sentence, evaluate the representation relative to the context of the experimental situation, and revise their representation when necessary. This revision process is precisely the kind that children have been shown to have difficulty with in comprehension. The key difference is that the ambiguity inherent in the NS stage is one that exists only for children up to the age of 3½ years, unlike attachment or scopal ambiguities, which are also ambiguous for adults. When children exit the NS stage, their grammar changes, and the processor is no longer faced with the problem of ambiguity resolution.

Note that these issues do not seem to impact language production in the same way, and indeed, whether a child in the NS stage produces an overt versus null subject seems to be sensitive to pragmatic-discourse factors, including givenness, person, and animacy (Hughes & Allen, 2008).

Overall, the experimental findings demonstrate that children in the NS stage have a non-target grammar. Additionally, while performance factors are not the underlying cause of the NS stage, they appear to play an important role in the resolution of the mood ambiguity associated with NS utterances. Children’s processing resources do not allow them to integrate all sources of information needed to resolve the ambiguity and this leads them to assign one single parse and then stick with it.

**Later language development: A-movement**

Once children have moved beyond telegraphic speech (including the RI and NS stages), their language development is incredibly rapid, and by the time they are five years old, they make very few errors in production. However, comprehension data shows that children still fail to assign a target representation to a small set of sentence structures. Do these comprehension errors reflect a non-target grammar, and if so, is it consistent with UG? Conversely, are the errors due to extra syntactic limitations, despite the children having acquired the adult grammar?

The development of sentences involving A(rgument)-movement, such as verbal passives (9), is arguably the most widely studied instance of this type of late acquisition. One prominent hypothesis – the A-Chain Deficit Hypothesis (ACDH) (Borer & Wexler 1987) -- holds that the
grammatical ability to form A-chain emerges maturationally in the child at around age 6 or 7. In A-movement, a DP is displaced from one argument position of a sentence to another (e.g. object to subject position in passives). English-acquiring children rarely produce the be-passive, in either naturalistic (e.g. Harwood, 1959) or elicited (e.g. Hayhurst, 1967) speech, although in some experimental conditions, children can be primed to produce verbal passives (Messenger et al. 2012; Bencini and Valian 2008).¹ English-acquiring children are also delayed in showing adult comprehension (Slobin, 1966) of both “short” passives, where the by-phrase is missing (9a), and “long” passives, where it is pronounced (9b) (Hirsch & Wexler, 2006, but see O’Brien, Grolla and Lillo-Martin 2006).² This non-adult comprehension is attested in many (though not all) languages, including Afro-Asiatic, Sino-Tibetan and Altaic languages as well as Indo-European ones. The most notable exception to the generalization is Sesotho, in which the verbal passive seems to be acquired early (Demuth, 1989; Demuth et al. 2010, but see Crawford 2009 for results suggesting a delay for a subset of passives in Sesotho. See also note 4).

(9) a. Ernie was seen.  
   b. Ernie was seen by Bert.  

However, it is clearly not the case that children are delayed in acquiring all A-movement, contra the predictions of the ACDH. For example, in the active voice, subjects undergo A-movement out of the verbal domain (Koopman & Sportiche, 1991), yet children have no difficulties correctly placing the subject outside the VP (Stromswold, 1996). Very young Italian and French-speaking children have acquired various kinds of reflexive structures (impersonals and medio-passives) which involve A-movement (Snyder & Hyams, to appear). Recent experimental work shows that children have Subject-to-Object Raising (10) from as young as 3 years old (e.g. Kirby, 2010).

(10) Ernie wants Bert [Bert to win the race].

Given this range of results, the acquisition of A-movement must be approached with the goal of understanding why some structures are acquired late and others early. Insight into this question is provided by the acquisition of another A-movement structure: Subject-to-Subject Raising (StSR). StSR sentences involving the verbs seem and appear are noted to be delayed in English and Dutch-acquiring children (Hirsch, Orfitelli, & Wexler, 2008).

In a series of seven experimental studies, Orfitelli (2012) compared the acquisition of StSR with seem (11) to those of other English StSR predicates, illustrated in (12).

(11) Ernie seems (to Bert) [to dance].

---

¹ Children’s primed passive responses include many get-passives, which are known to be acquired earlier, and argued to be structurally distinct from ‘canonical’ verbal be passives. Children’s responses also included many “reversed passives”, i.e., passive morphology but with the thematic mappings of the active version. In these cases children have a passive “syntactic frame” but fail to execute the movement and/or theta-role transfer operations, that are argued to be the locus of children’s delays.

² For various reasons, in some languages (including English) there is an effect of verb-type. Passives with ‘actional’ participles are acquired much earlier, around 3-4 years. Those with ‘non-actional’ participles are delayed in comprehension until later (Maratsos et. al 1985, among others). See Orfitelli 2012 for a discussion of these facts.
(10) Ernie is about/is going/tends (*to Bert) [to dance]

As in previous studies, children demonstrated non-adult comprehension of StSR with the verb *seem* (11) until approximately 6-years-old. Importantly, however, these same children exhibited adult comprehension of StSR sentences with *be about, be going, and tend* from as young as 4 years-old. Further, a within-subjects comparison of 30 children found an over 96% correspondence between the comprehension of verbal passives and *seem*-type StSR (Table 28.1). These results show that children have no difficulty with the process of StSR itself; rather, it is a specific property of StSR with *seem* and *appear* that leads to their delay. Second, whatever this property is, it is shared with the verbal passive.

Table 28.1: Number of children in Orfitelli 2012b showing above chance performance for StSR sentences with *seem* and *(be) about* as compared to non-actional passives.

<table>
<thead>
<tr>
<th></th>
<th>StSR</th>
<th>StSR</th>
<th>Short</th>
<th>Long</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(<em>seem</em>)</td>
<td>((be) about)</td>
<td>non-actional passive</td>
<td>non-actional passive</td>
</tr>
<tr>
<td>4 years (N = 10)</td>
<td>0</td>
<td>10</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>5 years (N = 10)</td>
<td>1</td>
<td>10</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>6 years (N = 10)</td>
<td>7</td>
<td>10</td>
<td>6</td>
<td>7</td>
</tr>
</tbody>
</table>

As with the previously discussed phenomena, accounts of A-movement acquisition can be divided into competence-based theories and performance-based theories. Although the major competence-based theories (Borer & Wexler, 1987; Babyonyshev et al., 2001; Hyams et al., 2006; Hyams & Snyder, 2005; Snyder & Hyams, to appear) differ in their exact implementation and predictions, they share the assumption that children’s grammar is not target-like. Here, we will focus on one grammatical account, the *Argument Intervention Hypothesis* (AIH), proposed by Orfitelli (2012):

(10) Children are delayed in acquiring exactly those structures which require A-movement across an intervening argument.

The AIH appeals to a difference between the two types of StSR predicates: *seem* and *appear* (11) optionally permit an additional experiencer argument that structurally intervenes between the base and final positions of the argument that undergoes raising (Collins, 2005a). This is in contrast to other raising predicates (12), which never permit an experiencer. Movement of the embedded subject argument to the matrix subject position therefore violates universal conditions on syntactic locality such as Relativized Minimality (Rizzi, 1990).

One of the benefits of such an analysis is that it straightforwardly predicts the developmental relationship between *seem*-type StSR and the verbal passive. Under analyses in which *by*-phrase of the English passive is initially merged in the same position as the subject of
active sentences (Collins, 2005b), it intervenes between the base and final position of the promoted object.

These acquisition data are also in line with the idea that there is a ‘covert’ by-phrase in the passive that is syntactically and/or semantically active even when unpronounced (Baker, Johnson, & Roberts, 1989). As noted above, children are equally delayed in comprehending verbal passives whether the by-phrase is explicit or implicit. Children are similarly delayed on seem-type StSR whether or not the experiencer is pronounced, suggesting that it is also covertly present.3

Beyond verbal passives and StSR, the AIH makes a clear theoretical prediction regarding the acquisition of A-movement cross-linguistically: children will be delayed in all structures that involve movement across an intervening argument, and not delayed in those which do not. Experimental data supports this conclusion. As noted, VP-internal subject movement, Romance reflexives, and Subject-to-Object Raising are mastered early, as predicted, because these structures do not involve movement over an intervening argument. Conversely, the inverse copula construction (14), which under Moro’s (1996) analysis involves promotion of a predical DP over an intervening subject DP, is not comprehended by English-acquiring children until 6 to 7 years old, the age at which they begin to understand StSR and verbal passives (Hirsch & Wexler, 2007).4

(13) The cause of the fight is the bully.

To conclude, there appears to be a growing body of evidence in favor of the hypothesis that intervention/minimality is in some sense an inviolable constraint for young children. This is distinct from adults, who have mechanisms to circumvent it in some cases, allowing them to interpret verbal passives and seem-type StSR (see Collins, 2005 for one suggestion and also Snyder and Hyams to appear for another). Indeed, the AIH and similar proposals could serve as a tool to adjudicate between multiple possible syntactic structures. For example, in the case of the middle voice (15), it is unclear whether the understood external argument—which cannot be overtly expressed—is syntactically present (Ackema & Schoorlemmer, 2006, and references therein). If indeed it is present, we predict a cross-linguistic delay in the acquisition of the middle voice, with similar timing as the delay in verbal passives, StSR, and inverse copulas.

(14) The toys sell easily.

Intervention effects in acquisition have been noted for A’-movement as well, although they seem to disappear around 4-5 years old. (Friedmann et al 2009; Belletti et al 2012). For the

3 A related alternative is proposed by Choe (2013), who suggests that children’s difficulty with intervention is a performance effect driven by limited processing resources that prevent them from constructing a dependency over an overt experiencer (e.g. Bert in (11)). Such an analysis cannot explain children’s equal delay on “short” raised sentences where the experience is unpronounced, as well as their commensurate delay on short passives where the by-phrase is silent.
4 Snyder & Hyams (to appear) offer an explanation, also in terms of minimality, for the apparent earlier acquisition of passives in the specific discourse contexts given in the O’Brien et al. 2006 experiments and also in languages such as Sesotho (Demuth et al. 2010).
current discussion, we will restrict ourselves to intervention in A-movement, although a link may exist between the two phenomena.

**Control**

The recursive embedding of one constituent inside another is the feature of human language that accounts for its infinite expressive power. For this reason the child’s acquisition of sentential embedding represents a giant developmental leap. In the previous section we discussed the development of one kind of embedded structure, involving raising. In this section we discuss another type of complex sentence: control structures, which are superficially identical to raising structures, but with very different syntactic and semantic properties.

In general, embedded sentences can be tensed (16a), or infinitival (16b). In a tensed clause all the grammatical functions associated with the verb are expressed, e.g. *Ernie* is the subject and *the piano* is the object of the embedded sentence in (16a). Most often, however, the subject of an embedded infinitive is not overtly expressed (16b). We refer to this “silent” subject as PRO.

(16)  a. Bert thinks that [Ernie plays the piano].
    b. Erniei likes/wants [PROi to play the piano].

Sentences like (16b) are referred to as control structures because the identity or reference of the embedded subject (PRO) is determined by the matrix (higher) subject (Ernie is the liker and the player). Because control sentences involve this silent material, they pose a potential learning challenge for the child. The problem is made apparent when we also consider the sentences in (17).

(17)  a. Bert told/persuaded Erniei [PROi to play the piano].
    b. Berti promised/threatened Ernie [PROi to play the piano].

In (17b) the matrix subject controls PRO; in (17a) the object, *Ernie*, is the controller (i.e. Ernie is the piano player). A possible generalization that emerges is: the structurally closest nominal controls PRO (object if there is one, otherwise the subject), but (17b) is a clear exception - *Bert* is the controller of PRO (i.e. Bert is the piano-player).

As in (16b), the embedded sentence in (17) is a complement to the matrix verb, that is, the higher verb requires a direct object and a sentential/propositional object. Control also occurs into temporal adjunct clauses, as in (18). In temporal (e.g. *before, after, while*) clauses PRO is always controlled by the matrix subject. This is assumed to be for structural reasons; the position at which the adjunct attaches to the matrix clause precludes the object from controlling PRO.

(18) Berti hugged Ernie before/after PROi playing the piano.

Not surprisingly, children take time to fully acquire an adult system of control. Beginning with the seminal work of C. Chomsky (1969), various experimental studies (Hsu et al. 1985; Cairns, McDaniel, Hsu, Ryan, & Rapp, 1994; Goodluck, 1981; McDaniel, Cairns, & Hsu, 1991; Wexler,
1992, Brohier & Wexler, 1995, among others) have shown that many children ages 4 to 5½ (though not all) go through similar stages in the development of control (19):

(19) (i) Children initially take PRO to have free reference (i.e. to refer to an independent figure in the experimental setting), in both complement (17a) and adjunct clauses (18).

(ii) Children persist in assigning free reference in adjunct clauses (18) after they have sorted out object control in complement structures (17a).

(iii) Difficulties with promise type verbs persist for a longer time than like or tell type verbs. As an example, for the 4-year old, the player in (17b) will generally be the object, Ernie.

(iv) Other children (or the same children at a later stage) seem to require that PRO in adverbials (18) be controlled, either by the subject of the matrix clause (adultlike) or the object (non-adultlike).

(v) These children also seem to extend the control relation to overt pronouns in sentences like (20). This has been referred to as the adverbial coreference requirement (McDaniel et al., 1991).

(20) Berti washed Erniej after hei/j/k swam in the pool.

So the child who requires PRO to have Bert as controller in (18) will also interpret he to refer uniquely to Bert in (20). This is a possible, though not required, meaning for adults, for whom he can refer to any male individual given an appropriate context.

The non-adult interpretations noted in (19) have been found in various experiments using both act-out and/or judgment tasks. These results led Hsu and colleagues to propose a stage model of control, according to which a child might construct a series of grammars (based on different attachment sites for the complement or adjunct clause) before arriving at the adult system.

Two sorts of analyses exist for why children have free reference of PRO. According to the stage model, children initially have a “flat” structure for control sentences, that is, the complement or adjunct clause is not subordinated to the matrix clause, but is more like a coordinate structure (Cairns et al. 1994; Goodluck, 1981; McDaniel et al. 1991). As a result PRO is not in the right structural configuration to get an adult control reading. McDaniel and Cairns (1990) suggest that children may resort to the flatter structure for processing reasons. Basically, the child treats the two “conjuncts” as independent clauses, an analysis that requires less computational resources because the child need not keep the matrix clause in working memory while analyzing the subordinate clause. Another prominent hypothesis is that children have not yet acquired PRO or the syntactic operation necessary for linking the event time of the adjunct to the event time of the matrix. Lacking the adult syntax they treat the embedded clause as more of a nominal, roughly as in (21) (Wexler, 1992; Carlson, 1990). In (21) anybody can be the piano player. This ‘nominalization hypothesis’ (originally suggested by Carlson, 1990) would account for the findings in (19i-ii).

(21) Bert hugged Ernie before (the) playing (of) the piano.
Brohier and Wexler (1995) question assumptions and empirical basis of the stage model. Based on a reexamination of previous results and their own experiment using a truth value judgment task, they find very few children with non-adult (i.e. obligatory object) control in adjuncts (18). Rather, they show that children either have adult (i.e. subject) control or they allow ‘free reference’ of PRO, which on their analysis arises because of the nominalization strategy.

While the nominalization hypothesis is consistent with the free interpretation of PRO, there is thus far little independent evidence for this analysis (but see Goodluck, 1991 for some weak support). For example, children are unlikely to spontaneously produce nominalizations, and experimental work suggests that children between the ages of 3½ and 5 (the same ages of the children in the control studies) do not have an adult interpretation of nominals like the *kicking of him*, choosing ‘him’ as the subject rather than the object of the action (Roeppe, 1978; de Villiers et al. 1995). Also, it is unclear what would force children to give up the nominalization structure in favor of an adult control analysis, given that the nominalization is consistent with any possible controller, viz. (21) is true whether Bert or Ernie or anyone else is playing the piano.

An alternative to the stage and nominalization hypotheses might relate children’s non-adult performance with control to their behavior with raising structures (discussed in the previous section). Perhaps the presence of an intervening argument (*Ernie*) in sentences like (17a) and (17b) blocks the connection between PRO and the matrix subject. As a result PRO remains “unlinked”, i.e. free in reference (property 19ii), unless it associates with the intervener itself, i.e., is object controlled. This kind of analysis, based on the argument intervention hypothesis (AIH) would also account for why promise type verbs are especially difficult (see also Snyder & Hyams, to appear). In this case, the child must also skip over an intervening argument to establish the adult (subject) control relation. Lacking the ability to do so, he associates PRO with the intervener itself.5 The intervention explanation is particularly natural under a newer ‘movement theory of control’ (MTC) (Hornstein, 1999, 2003), according to which “control” is derived by movement of the embedded subject to a matrix argument position (e.g. Ernie in (23b), similar (though not identical) to raising (22a). According the MTC, control into adverbial clauses involves “sideways” movement out of the adjunct. The marked status of this operation might explain why children show a more delayed acquisition of control into adjuncts relative to complement clauses. Given the MTC and other recent advances in the syntactic analysis of control (e.g. Landau, 2000), the acquisition issues surrounding control seem ripe for further study.

**Control versus raising**

As noted previously, control structures are superficially similar to raising structures. The two sentence types are repeated in (22).

(22) a. Ernie seems [to play the piano]. raising

5 Free reference in complement clauses such as (17a) is more difficult to explain (property 19i) via the AIH. However, this stage is attested in far fewer subjects (e.g. 2 out of 20 children in McDaniel et al. (1991) and 9 observations out of 45 in Cairns et al.’s (1994) longitudinal study. It should be noted as well that in act-out tasks children allow free reference for PRO in adjuncts but not in complement clauses (Goodluck, 1981; Hsu et al., 1985).
b. Ernie likes [to play the piano]. control

It is reasonable to ask if the surface similarity of these two sentence types poses a learning problem for the child, a question that Becker (2006, 2009) explores in detail. How does the child know if he is dealing with a raising or control structure given that the embedded subject position is phonologically empty in both cases, an even more perspicuous problem if control reduces to movement (Hornstein 1999, 2003). In principle, expletive it could tell the children that they are faced with a raising verb and not a control verb, given the well known contrast in (23) (Postal, 1978). However, as noticed by Becker, a more child-friendly difference between the two verb types concerns the animacy of the subject. Verbs like seem can occur with inanimate subjects, while like-type verbs generally require animate subjects, as shown in (24).

(23)  a. It seems that Ernie plays the piano. (Ernie seems to play the piano)  
     b. *It likes that Ernie plays the piano. (cf. Ernie likes to play the piano)

(24)  a. The rock seems to be heavy/ The flower seems to open at dawn.
     b. *The rock likes to be heavy/ *The flower likes to open at dawn. 
     (cf. The dog likes to chase the ball)

Thus, probabilistic cues in the input such as (in)animacy (and others discussed by Becker) would allow the child to separate control and raising verbs despite their appearance in superficially similar structures. And given that the two structures have very different developmental profiles, it seems likely that they are indeed distinguished by children at a young age.

General conclusions

In this chapter, we have discussed four notable stages of language acquisition. Early in their sentence production, children optionally omit both verbal inflection (RI stage) and subjects (NS stage), and there is evidence from both production and comprehension that these omissions are due to a difference between the child and adult grammars. Later in development, children whose productive language is largely adult-like nonetheless continue to show difficulties in comprehension of certain A-movement and control structures. Once again, these errors seem to reflect the child’s developing grammar. In addition, each of these areas also reveals a role for extra-syntactic - processing and /or pragmatic - factors in language development.

Universal Grammar is a theory about the form of grammar the human mind can acquire under normal conditions of language exposure and cognitive growth, and in this sense it also defines the grammars a child can entertain in the course of development. And while linguistic theory does not tell us directly why the 2-year old opts for a NS or RI grammar, or why 5-year olds adhere more strictly to intervention than their parents, it does provide a framework for understanding these stages. Where children’s output does not reflect their input, as in the cases discussed here, we might assume that there is some cognitive pay-off in this option, either on a domain general or domain specific level. Perhaps they have settled (temporarily) on a system that is computationally simpler, either in terms of linguistic representation or processing, or both.
The goal of a comprehensive theory of language development is to understand not only the various stages and their underlying grammars, but also to explain why the child takes a particular path to the target grammar. In large measure this will require a deeper understanding of what is “complex” for the child – whether defined by a theory of linguistic performance or by grammatical theory more directly.

References


Swinney, David, and Penny Prather. 1989. On the comprehension of lexical ambiguity by young children:


