The Acquisition of Raising

by

Christopher K. Hirsch

B.A. Cognitive Science
B.A. Linguistics
University of California, Berkeley
2002

SUBMITTED TO THE DEPARTMENT OF COGNITIVE SCIENCES IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE DEGREE OF

DOCTOR OF PHILOSOPHY IN COGNITIVE SCIENCES
AT THE
MASSACHUSETTS INSTITUTE OF TECHNOLOGY

FEBRUARY 2011

Signature of Author: ______________________
Department of Brain and Cognitive Sciences
October 27, 2010

Certified by: ______________________
Kenneth N. Wexler, Ph.D.
Professor of Psychology and Linguistics
Thesis Supervisor

Accepted by: ______________________
Earl K. Miller, Ph.D.
Picower Professor of Neuroscience
Director, BCS Graduate Program

The Acquisition of Raising

by

Christopher K. Hirsch

Submitted to the Department of Brain and Cognitive Sciences on October 27th, 2010 in Partial Fulfillment of the Requirements for the Degree of Doctor of Philosophy in Cognitive Sciences

Abstract

This work serves as the first comprehensive investigation into typically developing children's acquisition of subject-to-subject (StS) raising. In particular, it asks how English-speaking children come to comprehend a StS raised sentence like (1) versus its semantically equivalent unraised counterpart (2), and how the presence of an experiencer-phrase affects interpretation (3):

(1) John_i seems [t_i to be dancing].
(2) It seems that John is dancing.
(3) John_i seems to Mary [t_i to be dancing].

The acquisition of StS raising is of particular bearing given renewed interest in the acquisition of verbal passives, which share syntactic traits with StS raising and which have been argued to develop late under genetic guidance.

Using sentence-picture matching and truth-value judgment experiments, the following comprehension results obtain:

- Unraised sentences are acquired early, with most three-year-old children demonstrating mastery.
- Sentences involving StS raising over an experiencer (ROE) are delayed until around age seven, with many children incorrectly interpreting them as involving raising-to-object (RtO) syntax.
- Sentences involving StS raising with no experiencer (RNE) are likewise delayed until around age seven, with many children incorrectly interpreting them as involving either subject control or copular syntax.
- Subject control sentences are acquired early, by at least age three.
- The delayed acquisition of StS raising (ROE and RNE) appears to be developmentally linked with the delayed acquisition of verbal passives.

The noted delay in comprehension of StS raising occurs despite the fact that StS raising sentences are found to be relatively common in child-directed speech. These data

serve to rule out several grammatical acquisition accounts in the literature that have attempted to capture children's delayed comprehension of verbal passives (e.g. External Argument Requirement Hypothesis, Canonical Alignment Hypothesis, and Universal Freezing Hypothesis). The data, however, are both compatible with and predicted by the Universal Phase Requirement (UPR; Wexler, 2004), a grammatical account on which children are claimed to take all *v*Ps to define strong phases for maturational reasons. Only UPR correctly predicts delay for StS raising (with and without an experiencer-phrase) in early child grammar, and a positive correlation with the acquisition of verbal passives.

To Jennifer. For patience, near limitless.

ACKNOWLEDGMENTS

Immense gratitude to Ken Wexler for many years of inspiration, long talks, an open door, insight, cheerleading, and first class mentorship.

Special thanks to the rest of my committee, Shigeru Miyagawa, Irene Heim, and Laura Schultz, for insightful comments, suggestions, and critiques.

I was blessed with very loving parents, Steven Hirsch and Susanne Hirsch; early lessons learned helped drive the passion that underwrote this work. Mom also served as my (saintly) editor.

Of everyone, I am the one most thankful for finishing. My wife, Jennifer Hirsch, comes in second. A supportive spouse makes nearly anything seem possible.

Two longtime lab peers, Alexandra Perovic and Nadya Modyanova, deserve particular recognition. It is in the trenches of daily research that the strongest academic bonds are formed.

Many others along the way offered invaluable assistance: David Pesetsky, Danny Fox, Ted Gibson, Yosef Grodzinsky, Alec Marantz, Gregory Marton, Misha Becker, Nina Hyams, Carson Schütze, William Snyder, Karen Froud, Vina Tsakali, Jenny Ganger, Loes Koring, Sophia Manika, Charlotte Herzmann, and the audiences at various conferences.

Strong experimental research is made easier through a strong laboratory. I am thankful to the entire Wexler ab/Normal Language Lab, and in particular Sun Yoo, Jeremy Hartman, Karen Law, Robyn Orfitelli, and the rest of the Wexler Lab undergraduate research assistants.

Money always helps. This research was supported in part by an NSF Graduate Fellowship, and by a Marcus Fund awarded to the Department of Brain and Cognitive Sciences at MIT.

The work presented here requires participating children. I am grateful to all the parents and teachers at the various daycares for their consent, patience, and willingness to support these scientific inquiries. Most importantly, thank you to all the children who actually had to sit through and endure all these various experiments.

Table of Contents

Chapter 1: Grammar Acquisition

1.1 Learning and Maturation

Young children are remarkable in their ability to quickly acquire new information, and one cognitive domain where this is readily apparent is that of first language acquisition. Humans are born into the world unable to produce any meaningful words (never mind sentences), nor do they have the capacity to comprehend the speech of the adults around them. Yet, within just a few short years, they are already producing multi-word utterances and responding intelligently to statements directed to them (Boysson-Bardies, 1999). Furthermore, from the time of their first utterances, they demonstrate a deep understanding about the specific syntactic rules that govern their language (Pierce, 1992a, Poeppel and Wexler, 1993).

Such early linguistic knowledge is all the more impressive given that children basically only draw from positive evidence (grammatical sentences to which they are exposed), receive extremely little negative evidence (about which sentences are ungrammatical), and in very short order come to a capacity for the comprehension and production of a potentially infinite number of novel sentences, having been exposed only to a finite number of them (for more on the logical problem of language acquisition, see Baker and McCarthy, 1981).

A central approach to dealing with the question of how children are able to so quickly figure out the details of the language of their environment is the Innateness Hypothesis (IH; Chomsky, 1959). On this Nativist view, the solution is that a substantial amount of linguistic knowledge is inborn and therefore need not be learned. Those constraints that hold universally in the world's languages are assumed not to be taught or learned, but are taken to be innate (with the strong prediction that any such aspects will be subject to a similar time course of development). IH obviously does not posit that *all* linguistic knowledge is innate (since children in different linguistic environments acquire different languages), but conjectures that variation in human language is not unlimited, and that Universal Grammar (UG), the part of language taken to be innate, defines the range of possible variation. Language acquisition, then, is what results from an interaction of the inborn factors and the extrinsic environmental ones. UG cannot account for the lexicon, regularities in a language, or how innate constructs are instantiated in their linguistic environment (Fodor, 1966). Those linguistic aspects must all be determined through learning mechanisms.

UG itself is assumed to consist of two types of constraints, namely principles and parameters (Chomsky, 1981). The principles encode those properties of languages that are universal, that is, that are shared across all languages. Under the assumption that all humans share the same basic genome, we expect those aspects of language that are universal to be among those that are genetically determined. The parameters code for the linguistic properties that vary from language to language. Those areas showing great crosslinguistic variation are less likely to have been innately specified (i.e. they are learned). It is assumed, though, that both the principles and parameters are specified by UG (i.e. innately determined). The child must set, on the basis of experience (i.e. linguistic input), the parameters to the correct values for his particular linguistic environment. Indeed, it is well known that children are brilliant learners of language-particular phenomena (i.e. parameter setting) (Wexler, 1998). Anyone who has ever tried carrying on a conversation with a typical five-year-old is acutely aware of how "adult-like" they sound grammatically. Certainly their vocabulary, and depth and breadth of subject matter, will not be at the same level as an adult's, but for the most part, their grammar is quite developed (see Section 1.2.3 for discussion of some of these aspects of syntax that emerge early).

Given evidence that *something* (UG) is built-in (innate), two possibilities present themselves. First, the process of language acquisition can be conceptualized as a child learning to set parameters on the basis of linguistic exposure, where the principles and parameters themselves are present in child grammar from the start. This idea that what is built-in is there from the start and does not change is known as the Continuity Hypothesis (CH; Pinker, 1984). This has generally been the narrow view of language acquisition in the generative approach to the discipline. On this theory, child grammar *is* adult grammar. All that differs is that adults have learned to set the relevant parameters correctly. Certainly learning of this nature (knowledge reached through experience) is crucial to the goal of reaching adult language competency.

What CH seems to ignore, however, is the role that biological maturation might play in making both the principles and parameters available to the young child. Just because a capacity is innately endowed does not mean it is expressed from birth. Certain biological machinery must be in place in order to achieve certain physical and cognitive outcomes.[1] An alternative to CH therefore is the Maturation Hypothesis (MH), which

[1] This is a rather crude example, but serves to make the point. Healthy human females can reproduce sexually. Yet, no matter what, a seven year-old girl is not going to get pregnant and bear children. That does not mean the biological machinery needed to support pregnancy/reproduction is not innately specified

posits that grammatical principles are innately specified, but may take time to develop under genetic guidance (Borer and Wexler, 1987). There is no conceptual reason to suppose that UG itself is not subject to development, and as such, that certain principles and parameters might develop late under genetic guidance. On MH, certain grammatical representations allowed by UG are ungrammatical for young children because of constraints imposed by their particular immature biology as opposed to adult biology. This immaturity is usually assumed to be related to yet undeveloped neural substrates required for the representation and processing of the relevant grammatical structures.

An obvious challenge for CH is the apparent fact that child grammar often does differ significantly from adult grammar. According to CH, any such differences must be due to either failure to have correctly set certain parameters or to extra-linguistic considerations. On MH, in addition to those two possibilities, there also exists the prospect that delays in child grammar reflect a cognitive state in which relevant linguistic operations and structures have not been made available due to biological (i.e. neurological) immaturity. For aspects of grammar that are late to be acquired, it must be asked: Is the delay due to a deficiency in input (i.e. learning), to prematurity (i.e. maturation), or to other extrinsic factors (e.g. inattention, processing difficulties, experimental weakness, pragmatic miscues, etc)?

In particular, it must be asked why some structures are delayed even until late into childhood, especially when there is evidence that a particular structure is abundant in the input. Why should it take so long for some structures to develop? It is precisely these cases that seem to defy learning explanations (given children are so adept at parameter

---

(it mostly certainly is), but rather, that such biological processes are not present until development (i.e. puberty) has taken place under genetic guidance.

setting). Seeking evidence for MH might therefore offer solutions. If evidence is found that some aspects of grammar develop late for maturational reasons, it opens the door for new methods of doing neuroscience. One can envision attempting to correlate the time course of development of a linguistic process that is known to emerge under genetic (biological) guidance with what neural structures develop at that time, and use that correlation to probe for possible causation. With potential neural language substrates in hand, one can then look for genetic markers known to influence development of such brain structures, as well as examine what happens in adults who suffer injury or develop diseases that affect those brain regions. Such convergent methodologies might some day help answer foundational questions about how the brain instantiates language, and about what areas of the brain are both *involved* and *crucial* for language comprehension.

The study of first language acquisition also speaks to other general issues in the cognitive sciences. At least in the generative approach, language development research makes great use of linguistic theory (for an excellent review, see Guasti, 2002). This is a particular positive attribute of the science that can be done in the field of language acquisition. In many areas of cognitive science, the cognitive models themselves are rather sparse. This is not surprising given the difficult task of gathering the necessary empirical data to decide between competing models. Linguistic theory on the other hand provides remarkably rich cognitive models, due in good part to the easy access to abundant data (any native speaker can make thousands of linguistic judgments). By tapping into these richly articulated linguistic theories, research into language acquisition gains a very deep and strong modeling foundation. Ultimately, though, it is hoped that the study of language development can work bi-directionally to contribute back to inform linguistic theory. Furthermore, by examining the time course of language development, we can integrate the study of language into the broader study of biological development. The hope persists that this type of analysis will play a role in the genetic underpinnings of language, as it has already done in some areas of grammar.[2]

One area of grammar notorious for demonstrating late development, and which might bear on MH, involves various kinds of long-distance dependencies. In this dissertation, evidence is explored for the late development of one syntactic structure in particular: subject-to-subject raising. It is hoped that such an investigation will in due course lead to evidence for the role of biology in the development of linguistic structure.

1.2 Verbal Passive Acquisition

A review of the syntax acquisition literature as it applies to typically developing children highlights several interesting areas for helping to tease apart the competing aspects of learning versus maturation, and sheds light on many important questions in the cognitive sciences. One syntactic structure in particular that has received substantial investigation from language acquisitionists is that of the verbal passive. Before delving into what is known concerning the acquisition of such sentences, we turn to a few details concerning the syntactic analysis of verbal passives in adult linguistic theory.

[2] In particular, the development of obligatory finiteness is genetically determined, as argued in Wexler (2002). The latest and most systematic behavioral genetic evidence strongly confirms that the development of finiteness is controlled by genetics and that the genetic source of finiteness is independent of the genetic source of phonological working memory (Bishop, Adams, and Norbury, 2006).

1.2.1 *Basic Syntactic Theory and Verbal Passives*

Many transitive verbs have both an active voice form (1) and a passive voice form (2). The two are (for all intents and purposes here) semantically equivalent. The passive form, in English, differs from its active counterpart in that the subject bears the thematic-role usually assigned to the object in the active voice, consists verbally of *be* and a passive participle (identical in form to the past participle), and optionally takes a *by*-phrase which bears the thematic-role usually assigned to the subject in the active voice.

(1) John loves Mary.

(2) Mary is loved (by John).

One popular syntactic approach to capturing the semantic relationship between actives and passives has focused on simplifying the lexicon by assuming that there is only one form of each (transitive) verb, which always assigns thematic-roles to canonical structural positions, and where differences in surface form are the result of syntactic operations that can re-order elements (for reviews, see Haegeman, 1994, Chomsky, 1995).

On such an account, there is only one verb *love* in the lexicon, whose subject is associated with an experiencer theta-role and whose object is associated with a theme theta-role. When building an active sentence around this verb, something like (3) is realized (abstracting away significant syntactic detail for the moment).

(3) $[_{IP}$ John $[_{I'}$ $[_{VP}$ $[_{V'}$ loves $[_{NP}$ Mary$]]]]]$

At the representational level at which verbal arguments are projected and theta-roles assigned, the passive form has a similar basic structure to the active, ignoring for the moment how the *by*-phrase comes to obtain the theta-role assigned to the subject in the active voice (4).

(4) [$_{IP}$ e [$_{I'}$ [$_{VP}$ [$_{V'}$ loves [$_{NP}$ Mary]] (by John)]]]

The representation in (4) indicates that the logical object (internal argument) is always base generated in the same position regardless of voice, that is, as the sister to V, from which it receives its theta-role. The subject position is e(mpty), since there is no subject argument, just the optionally expressed post-verbal *by*-phrase adjunct. For syntactic reasons the object *Mary* must "move" to subject position, which derives the passive in (5).[3] The NP that comes to appear in subject position of a passive sentence is actually the same internal argument as in an active sentence; it is thus a syntactically *derived* subject.

(5) [$_{IP}$ [$_{NP}$ Mary]$_i$ [$_{I'}$ is [$_{VP}$ [$_{V'}$ loved t_i] (by John)]]]

When an NP is moved, it leaves behind a trace (*t*), an invisible (i.e. null) copy of itself. The trace is connected to its antecedent via an *A*(rgument)-chain (and hence the co-indexation for expository reasons). The movement of an element form an *A*-position to

[3] The reason for this movement is not of immediate relevance, but it can be assumed to be for either case reasons (Chomsky, 1981) and/or the Extended Projection Principle (EPP; Chomsky, 1981).

another *A*-position (e.g. object-to-subject movement) is considered *A*-movement. The *A*-position in which the NP is base-generated is called the *foot* of the chain, while the NP's final landing site is called the *head* of the chain (always an *A*-position in cases of *A*-movement). Theta-roles are assigned to the foot of a chain, while case is assigned to the head of a chain.

While conceptual considerations for such movement in verbal passives are strong (i.e. maintaining linking rules greatly simplifies the lexicon), there is independent empirical evidence for such object-to-subject movement. First, note that the subject position in a verbal passive is non-thematic. This can be seen in the contrast between (6) and (7), where the latter passive allows an expletive (which unlike other NPs does not require a theta-role).

(6) John believes that Mary is nice.

(7) It is believed by John that Mary is nice.

Furthermore, when an idiomatic subject is moved under passivization, it retains its idiomatic meaning, as noted in (8) and (9).

(8) John believes that the cat was let out of the bag.

(9) The cat is believed (by John) to have been let out of the bag.

There is also direct evidence that passives involve derived subjects (underlying objects). For example, resultative phrases (RPs) denote the resulting state that an object achieves after some action. In English, RPs may only be predicated of direct objects (Levin and Rappaport Hovav, 1995). So, in a sentence like (10), the RP *flat* can only modify the object *nail*; it cannot be used to denote the resulting state of the subject *John* (i.e. the sentence can only mean that the nail became flat as a consequence of John hammering, not that John himself became flat due to his actions).

(10) John hammered the nail flat.

Unergative verbs are thus incompatible with RPs, so that in (11), *tired* can only have a manner interpretation and cannot be interpreted as a RP (i.e. the sentence means John danced while he was tired, not that John became tired due to dancing).

(11) *John danced tired.

It might therefore appear surprising that a verbal passive sentence, which has no apparent post-verbal object, allows a RP (12). This "exception", however, is immediately understood if the underlying representation of (12) involves *the nail* having been base generated in object position (13).

(12) The nail was hammered flat.

(13) $[_{IP}$ e $[_{I'}$ was $[_{VP}$ $[_{V'}$ hammered [the nail] flat.

In Japanese, a numeral quantifier must be adjacent to (in mutual c-command with) the noun it modifies (Miyagawa, 1989). Thus in (14) the quantifier modifies the adjacent object, and (15) is not grammatical where the quantifier modifies the non-adjacent subject.

(14)	Taro-ga	**hon-o**	**3-satu**	katta
	Taro NOM	book-ACC	3-cl	bought
	'Taro bought three books'			
(15)	***Gakusei-ga**	hon-o	**2-ri**	katta
	student-NOM	book	2-cl	bought
	'Two students bought a book.'			

In the passive sentence (16), however, the numeral quantifier is grammatical even though it has "floated" away from the subject it modifies. This example can be assimilated with the earlier examples if it is assumed that the floated quantifier associates with the trace of the subject in object position. Thus, a quantifier can only be separated from a derived subject, which is what is found in verbal passives.

(16)	**Kuruma-ga**	doroboo-ni	**3-dai**	nusum-are-ta.
	car-NOM	thief-by	3-cl	steal-PASS-Pst
	'Three cars were stolen by the thief.'			

Some properties common to *A*-movement, as seen in passives, include the movement being necessary, the landing site of movement being an empty space (i.e. a position to which no theta-role is assigned), the landing site being an *A*-position, the position from which the element is moved is one where no case is assigned, the moved element leaves behind a co-indexed trace, which forms an *A*-chain, the chain is assigned one theta-role, and is case marked only once, at its head. Not all NP-movement, however, involves *A*-movement. In some cases, a NP is moved to a non-*A* (i.e. *A')* position. Such *A'*-movement is what obtains when moving a *wh*-phrase. *Wh*-phrases (or *wh*-constituents) include *wh* (interrogative) pronouns (*who, whom, what, where, when, why, how*) and *wh* determiner phrases (headed by *which, whose*). In many languages (e.g. Romance, Germanic), these *wh*-phrases must be overtly fronted (moved to [Spec,CP]). It can be shown that even in languages without overt *wh*-movement, *wh*-phrases nonetheless move to [Spec,CP] covertly (e.g. Chinese, Japanese). In English, fronting is mandatory. *A'*-movement is distinguished from *A*-movement on th basis of the landing site of *A'*-movement not being an *A*-position, and as such, *A'*-movement is not to a position that assigns case, so *wh*-movement is not case driven (Haegeman, 1994, Baltin, 2000). In (17), the *wh*-phrase *who* began as a direct object (sister to V), but moves to [Spec, CP], as seen in (18).

(17) Who$_i$ does John love t_i?

(18) [$_{CP}$ Who$_i$ [$_{C}$ does [$_{IP}$ [$_{NP}$ John]$_j$ [$_{I'}$ t_j [$_{VP}$ [$_{V'}$ love t_i]]]]]

With enough linguistic theory under our belt, we may now turn to the acquisition of verbal passives.

1.2.2 *Passives are Delayed in Acquisition (Early Evidence)*

It has been known for nearly a half century that verbal passives are rare in child-produced speech relative to adult use (Harwood, 1959, Wells, 1979). They are even rare in children's experimentally elicited speech (Maratsos and Ambramovitch, 1975, Horgan, 1978, Crain, Thornton, and Murasugi, 1987, Marchman, Bates, Burkardt, and Good, 1991). Of perhaps even greater interest are findings from comprehension studies that demonstrate young children have difficulties correctly interpreting verbal passives (Slobin, 1966, Sinclair, Sinclair, and de Marcellus, 1967, Turner and Rommetveit, 1967, Bever, 1970). In general, most of these studies find that children simply do not comprehend verbal passives, not that they interpret them as active voice transitive sentences.[4] So for example, if presented with two pictures, one showing a boy kissing a girl, the other depicting a girl kissing a boy, if asked which picture shows *the boy kissed the girl*, young children will correctly choose the picture with the boy kissing the girl, but if asked which picture shows *the girl was kissed by the boy*, very young children are likely to choose randomly (50% chance) between the two pictures.

These findings are made all the more interesting by several studies that in addition to noting a main effect of voice (passives are understood less well than actives), have all found an interaction of voice and verb type (where passives with some verb types are comprehended less well than with verbs of another class). The relevant verb classes are actional verbs, basically verbs with agentive subjects in the active voice (e.g. *comb, push, kiss, kick*), and nonactional subject experiencer verbs, basically those with experiencer

[4] Some studies with very young children (two to three years-old) hint at a stage where children might in fact interpret verbal passives as active voice sentences. In general, however, this is not what is found for errors with passives in older children. In any case, these particular details are not germane for the issues at hand. For discussion on these matters, see Hirsch and Wexler (2006c).

subjects in the active voice (e.g. *see, hear, love, remember*). In these studies, (English-speaking) children are significantly worse on passives with nonactional subject experiencer verbs compared to passives with actional verbs. Young children will generally have no difficulties correctly comprehending active voice sentences with either actional (19) or subject experiencer verbs (20). Yet when presented with semantically equivalent passive voice sentences, children are significantly worse on (22) compared to (21).

(19) The boy combs/pushes/kisses/tickles the girl.

(20) The boy sees/hears/loves/remembers the girl.

(21) The girl is combed/pushed/kissed/tickled by the boy.

(22) The girl is seen/heard/loved/remembered by the boy.

The findings with the active sentences, which use the same experimental materials as the passive sentences, demonstrate that children have no difficulty with subject experiencer verbs *per se*.[5] It is only in the passive voice that the effect of verb type is manifested. This basic finding has been found in every study that has crossed voice and verb type (Maratsos and Ambramovitch, 1975, Maratsos, Fox, Becker, and Chalkley, 1985, Gordon and Chafetz, 1990, Fox and Grodzinsky, 1998, Hirsch and Wexler, 2006a, Hirsch and Wexler, 2006c). It is apparent at basically all ages until around seven years,

[5] In fact, in Hirsch and Wexler (2006c), who investigated comprehension of both actional and subject experiencer verbs in active and passive sentences in 140 children, the authors actually found a slight, though statistically insignificant, performance bias for subject experiencer actives over actional actives. Children comprehend these verbs (in the active voice).

when children begin to successfully comprehend passives of both verb classes. This verb type passive effect can clearly be seen in Figure 1.2.1.

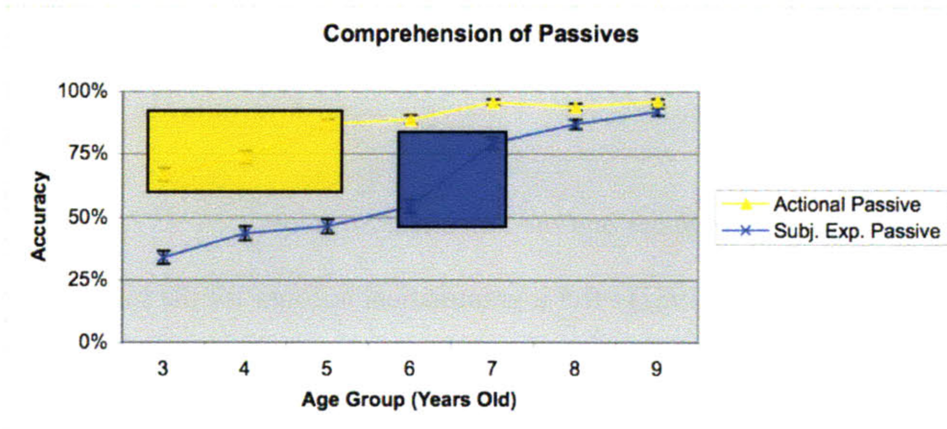


Figure 1.2.1: Comprehension of actional and subject experiencer verbal passives as function of age. Chart taken from Hirsch, Modyanova, and Wexler (2006).

What Figure 1.2.1 above makes clear is that there are two crucial periods with respect to passive acquisition. The first is indicated by the yellow box, and denotes the age range where actional passives are acquired. While some of the youngest children (three- to four-years-olds) have trouble with actional passives (and subject experiencer passives), most children appear to successfully comprehend actional passives by around age five. The second interesting age range involves just slightly older children. As already noted, these six- and seven-year-olds typically comprehend actional passives. As indicated by the blue box, this range from six to seven years-old is when most children acquire subject experiencer passives. Nearly all eight- and nine-year-olds comprehend both actional and subject experiencer passives, so this age range is of little interest to questions concerning acquisition (apart form establishing when ceiling performance is first reached). What accounts for these two jumps in performance? Do they involve the same developmental mechanism, and if so, why the age difference? What roles do learning and maturation have in explaining what occurs inside the yellow and blue boxes?

1.2.3 *A-Chain Deficit Hypothesis and the Adjectival Strategy*

Borer and Wexler (1987) offer a grammatical account for the interaction of voice (active vs. passive) and verb type (actional vs. subject experiencer). Their theory consists of two mechanisms working in conjunction to derive the interaction. First, a grammatical deficit is meant to capture the comprehension difficulties with verbal passives. Second, a linguistic strategy is posited to deal with why many young children are nonetheless able to comprehend passives of actional (i.e. agentive) verbs before passives with subject experiencer verbs, in spite of the hypothesized grammatical delay. They note that the grammatical difficulty is not a general difficulty with either all long-distance dependencies or general non-canonicity of thematic arguments. Empirical evidence from numerous acquisition studies in many languages would appear to substantiate this claim.

At ages where children demonstrate delayed knowledge of verbal passives, they do not exhibit parallel difficulties with structures involving *A'*-movement. Object-extracted *wh*-question are produced and comprehended at very young ages (Stromswold, 1995, Guasti, 2002). Importantly, no similar interaction is found for *wh*-questions (subject-extracted vs. object-extracted) and verb type (actional vs. subject experiencer), with three-year-old children comprehending all such *wh*-questions (Hirsch and Hartman, 2006a). Once pragmatic constraints on the licensing of restrictive relative clauses are satisfied, object-extracted relative clauses are found to be comprehended early (Hamburger and Crain, 1982, Hirsch and Wexler, In preparation). Topicalization of objects in V2 languages (e.g. German) is known to occur in even the earliest natural child productions (Poeppel and Wexler, 1993, Santelmann, 1995). Local object-scrambling, as found in Japanese, which has been argued to involve *A*'-movement (Tada, 1993), is also revealed to be early in acquisition (Otsu, 1994). Furthermore, children demonstrate no problems with Principle A of the Binding Theory, that is, the binding of reflexives (Wexler and Chien, 1985, Chien and Wexler, 1990), demonstrating that any difficulty with verbal passives is not simply a grammatical difficulty with discontinuous elements.

Borer and Wexler hypothesize that children's difficulties with verbal passives stem from their inability to form the *A*(rgument)-chain between the underlying object and its surface subject position (see Section 1.2.1). Due to the absence of the *A*-chains, the child has no means of establishing the thematic-role of the displaced object. According to Borer and Wexler, their proposed *A*-chain Deficit Hypothesis (ACDH) follows from human biology. Some biological (i.e. neural) structure is taken to mediate the cognitive representation of *A*-chains, and it is conjectured that this particular biological substrate matures late under genetic-guidance. While the psychological existence of *A*-chains is assumed to be innate (part of UG), their representation is hypothesized on ACDH not to be manifested at birth. Only after children's brains mature do *A*-chains become grammatical. Acquisition studies reporting delayed knowledge of verbal passives are thus taken to reflect children in the premature state, that is, subject to ACDH.

If a grammatical theory such as ACDH obtains to rule verbal passives ungrammatical for young children, some further explanation is needed to account for children's better comprehension of verbal passives involving actional verbs compared with subject experiencer verbs. For Borer and Wexler, the noted asymmetry is due to children's ability to apply a particular, learned linguistic analysis to actional passives, thereby deriving a grammatical parse that, while not exactly that of true verbal passives, is nonetheless "good enough" for children to respond in an appropriate fashion on those experimental paradigms that have been employed in testing passive comprehension. This particular analysis happens to be unavailable for subject experiencer passives due to linguistic constraints. Even without the syntactic (i.e. biological) means to represent *A*-chains, and therefore verbal passives, children nonetheless attempt to parse passives in their input. Borer and Wexler conjecture that English-speaking children actually analyze verbal passives (23) as (homophonous) adjectival passives (24).

(23) [The door]$_i$ was closed t_i

(24) The door was [$_{ADJ}$ closed]

While (23) and (24) have the same phonological surface form, they are taken by Borer and Wexler to involve different syntactic structures and derivations. Importantly, adjectival passives are taken not to contain the crucial *A*-chain that is present in verbal passives (following Wasow, 1977).[6,7]

Crucially for Borer and Wexler, the participles of actional verbs (25), but not the participles of (stative) subject experiencer verbs (26), allow licit adjectival forms.[8]

(25)	the doll appears combed; the combed doll	(Borer and Wexler, 1987, 7a)
(26)	*the doll appears seen; *the seen doll	(Borer and Wexler, 1987, 6a)

Since subject experiencer verbal passives do not have corresponding adjectival passives, children are left without any grammatical parse for passives involving such verbs. Thus while ACDH predicts equal delay for all verbal passives, the addition of the "adjectival strategy" deduces the comprehension interaction of voice and verb type, since actional verbal passives, but not subject experiencer verbal passives, may be analyzed as grammatical adjectival passives. With actional verbs, the adjectival passive serves as what Babyonyshev, Ganger, Pesetsky, and Wexler (2001) call a s(yntactic)-homophone for the verbal passive, a phrase with distinct grammatical structure, but common

[6] A similar conclusion, that children's early passive productions are adjectival in nature, appears to have been put forth by Watt (1970). Watt, however, does not go so far as to claim children interpret all passives adjectivally.

[7] Bresnan (1982) and Levin and Rappaport (1986) argue that verbal passive participles serve as the input to a lexical rule that derives the relevant adjectival passives. If these accounts are correct, such that verbal passives (which require *A*-chains) do indeed serve to feed the formation of adjectival passives, it would follow that verbal passives are a pre-requisite for the existence of adjectival passives, which poses a serious challenge to Borer and Wexler's story. Embick (2003) provides compelling arguments against the idea that verbal passives somehow feed the formation of adjectival passives, thus removing a possible challenge for those acquisition theories of passives involving both a grammatical deficit and an independent adjectival strategy that itself must not be subject to the relevant grammatical deficit.

[8] Obviously an adjectival strategy needs to account for the fact that many "actional" verbs make questionable adjectival passives (e.g. ?*the ball is hit*) and some "nonactional" verbs make fine adjectival passives (*the ball remains unseen*). These issues are outside the purview of this dissertation, but they form the basis for the theoretical and experimental investigation and subsequent claims in Hirsch and Hartman (2006a).

pronunciation. Comprehension of subject experiencer passives, which have no s-homophone, is predicted to only occur once biological changes unfold to license *A*-chains and the subsequent adult-like grammatical analysis of all verbal passives.

What is both exciting and maddening about Borer and Wexler's account is the boldness and novelty of the approach given the relative paucity of supporting data. That there might be two mechanisms at work (a grammatical issue and a compensatory strategy) seems rather innocuous (and quite reasonable). What is particularly daring, however, is the claim that the grammatical issue stems from maturational delay. This claim is ultimately based on a little more than a conceptual argument (i.e. if children's grammar appears at odds with UG, CH should be abandoned, and UG assumed to be subject to maturation). Unfortunately, Borer and Wexler do not offer a larger research program for better understanding what a maturation grammatical theory involves, entails, or predicts. Indeed, this topic has received very little discussion in the acquisition field. Rather, maturation is often just assumed or rejected based on the predictiveness of the grammatical account (in this case, ACDH). Little discussion centers on the concept of maturation itself. What is truly most astonishing, then, is that given recent evidence, Borer and Wexler appear to be generally correct about verbal passives being delayed due to biological immaturity (Section 1.2.5). First, however, some problems for their specific presumed grammatical deficit (i.e. ACDH).

1.2.4 *Challenges to ACDH*

At the time Borer and Wexler (1987) formulated ACDH, subjects were not assumed to be generated within the VP, but were assumed to be base-generated in subject position ([Spec, IP]). Borer and Wexler (1992) note that the VP-internal subject hypothesis (VPISH; Koopman and Sportiche, 1991) would prove difficult for their conjecture if movement of the VP-internal subject to its surface position involves the formation of an *A*-chain. Evidence arguing that children do not fail to raise subjects (Stromswold, 1996), strongly suggests that ACDH in its strong form (that all *A*-chains are ungrammatical) cannot be maintained. Note that ACDH faces a problem even on more recent accounts that completely dissociate the external argument from the VP, such that the subject is base generated in Spec, *v*P which is outside the VP (Kratzer, 1996). In this case, as with VPISH, the subject moves from one *A*-position to another *A*-position, thus forming an *A*-chain, which ACDH should rule out as ungrammatical.

A second empirical problem for ACDH comes from recent work in the syntax of adjectival passives. Unlike Wasow's (1977) lexical account, more modern accounts assume that adjectival passives are, like verbal passives, also derived syntactically (Kratzer, 2000, Anagnostopoulou, 2003, Embick, 2004). As such, they too involve an *A*-chain, and should therefore be ungrammatical on ACDH. This, however, would undermine the entire adjectival strategy which is supposed to explain the verb type effect in passive acquisition.

1.2.5 *Recent Evidence Supporting a Maturational Delay of Verbal Passives*

The heart of the Borer and Wexler proposal is a bipartite explanation for understanding the particular patterns of passive acquisition data, namely, the interaction of voice and verb type. The two necessary mechanisms are first a grammatical deficit (in their eyes, ACDH) and second a compensatory heuristic (their adjectival strategy), the former which is posited to always hold of premature children (for maturational reasons), while the latter applies asymmetrically depending on syntactic and semantic properties of the verbal participle and children's exposure to such factors (for learning reasons).[9] Three broad classes of evidence can be brought to bear on the validity of such a proposal. First, Borer and Wexler predict that in cases where no (adjectival) strategy obtains, verbal passives should only be comprehended once biological maturation takes place. Second, they predict evidence should exist supporting distinct mechanisms for the acquisition of those verbal passives with and without an s-homophone. Third, there is the testable prediction of the adjectival strategy, specifically, that before maturation takes place and thus prior to children's subsequent ability to comprehend passives with non-s-homophones (e.g. subject experiencer passives), children's early passives should be adjectival. Recently, evidence for each of these predictions has emerged.

Evidence for the maturational development of verbal passives is extensively discussed in recent work by Hirsch and Wexler (2006c), a selection of which is reviewed here. First, increasing cross-linguistic evidence is coming around to support the claim that true verbal passives are universally delayed in all languages.[10] In addition to the English data already discussed, delayed passive comprehension has been noted for a wide variety of languages, from a wide variety of language families: German (Grimm, Schöler, and Wintermantel, 1975, Bartke, 2004), Danish (Diderichsen, 2001), Dutch (Verrips, 1996),

[9] The precise detailing of what constitutes the relevant syntactic and semantic constraints governing children's adjectival strategy is outside the purview of this dissertation, but it is an active area of research in language acquisition (e.g. see Hirsch and Wexler, 2006a, Hirsch and Hartman, 2006b, and other works in progress).

[10] What constitutes a "true verbal passive" for the purposes of discussion is nothing more than whatever the relevant grammatical acquisition theory deems to be the relevant syntactic operation subject to maturation. For Borer and Wexler (1987), this was *A*-chains. As discussed in Section 1.2.6, the relevant grammatical operation is different for different acquisition theories. The use of the label "passive" is nothing more than a placeholder for actual grammatical operations. All of the relevant generative acquisition theories concern themselves with the maturation of operations/structures, not labels.

French (Sinclair, Sinclair, and de Marcellus, 1967), Spanish (Pierce, 1992b), (Brazilian) Portuguese (Gabriel, 2001, see Menuzzi, 2001 for discussion), Russian (Babyonshev and Brun, 2004), Serbian (Stojanovic, 2004, Djurkovic, 2005), Greek (Terzi and Wexler, 2002), Hebrew (Berman, 1985), Japanese (Sugisaki, 1998, Sano, 2000), and (Mandarin) Chinese (Chang, 1986). Indeed, in every language wherein actual comprehension experiments have been conducted, children have been shown to have trouble comprehending verbal passives.

Prior claims of "early" passive acquisition are coming under increasing scrutiny. Without exception, claims of languages purported to demonstrate early acquisition of passives are based solely on natural production studies. While production data can be a very valuable tool in carving acquisition theories, it is also limiting in that it often leaves an unclear picture as to children's underlying syntactic and semantic interpretations of the relevant productions. Where production data was taken as evidence for early passive acquisition in Kiche Maya (Pye and Poz, 1988), subsequent experimental studies demonstrate that contrary to the claims based on child-productions, children acquiring Kiche Maya do not comprehend verbal passives (Pye, 1992). In addition to overturning claims of early passive acquisition based on follow up comprehension studies, other such claims have been overturned on the basis of more detailed work at the theoretical syntax level. Another well-known argument for early passive acquisition comes from child-productions in Sesotho (Demuth, 1989). Crawford (2005) highlights various challenges for taking Sesotho-learning children to have mastered verbal passives. In reanalyzing the original Sesotho transcripts, Crawford finds that of the reported child-produced "passives", 22% are impersonal passives, which Demuth (1989) argues do not involve verbal passive (i.e. *A*-chain) syntax. Furthermore, approximately 30% of the reported passives are lexicalized (i.e. non-syntactic) forms, which can reasonably be expected to also not involve *A*-chains. Crawford further notes that Sesotho children fail to show voice alternation for most individual verbs, undermining the notion that they have a productive rule capable of translating between active and passive forms. Finally, Crawford offers a reasonable alternative non-passive syntactic analysis of the relevant child productions, arguing they are adversity passives with applicative syntax, which do not involve *A*-movement. Ultimately, the best counterargument to the claims of early passive acquisition in Sesotho comes from a recent comprehension study that found unambiguously verbal Sesotho passives to be delayed (Crawford and Hirsch, 2008). Claims of early passive acquisition in a handful of little-studied languages warrant further theoretical syntactic analyses, and actual comprehension experiments before strong conclusions can be drawn one way or another. Meanwhile, comprehension studies conducted in over a dozen languages with rich syntactic literatures all demonstrate delayed comprehension of verbal passives.

Another means of approaching the validity of the Borer and Wexler model is to consider what predictors exist for the two verb classes of verbal passives considered so far. The adjectival strategy predicts that learning mechanisms are crucial for the acquisition of actional passives, as this strategy has many lexical (and thus input-dependent) properties. That is, on the basis of the input, children must deduce the existence of the adjectival strategy in their language, if it applies at all. Subject experiencer passives are not subject to any such adjectival strategy (since they do not form adjectival passives), but rather require biological maturation, and as such should not be dependent on environmental factors.

Hirsch, Modyanova, and Wexler (2006) examined 16 environmental factors (e.g. age of parents at child's birth, household income, highest level of parental education, hours read to child per week, child's birth weight, native English fluency of parents) as predictors of children's comprehension of actional and subject experiencer passives using a parental survey. Five of the 16 factors were found to be statistically significant independent predictors for actional passive acquisition (parental education, parents' age at child's birth, child's age at daycare enrollment, hours read to child per week, and child's birth weight). These are all factors known to play an important role in determining general intelligence and language development. No factors, however, were found to significantly predict subject experiencer passive acquisition. Using all four predictors for actional passives together fails to significantly predict which children comprehend passives with subject experiencer verbs. Going so far as to build a model using all 16 factors together also does not provide a significant prediction equation for subject experiencer passive acquisition. The only true predictor of whether a child knows subject experiencer passives is his age. Those children younger than age seven very rarely comprehend subject experiencer passives, while almost all children older than age seven comprehend such passives.

Ganger, Dunn, and Gordon (2004) offer compelling behavioral genetics evidence in favor of a maturational account of verbal passive acquisition. They administered a passive test of both actional and subject experiencer verbs to groups of identical (monozygotic; MZ) and fraternal (dizygotic; DZ) twins. Twin studies are useful for teasing apart environmental (i.e. learning) factors from heritability (i.e. biological) factors. When it comes to within-twin pair correlation of comprehension scores, MZ twins are no more similar in their comprehension of actional passives than DZ twins. On subject experiencer passives, however, MZ twins show a much greater (and statistically significant) correlation than DZ twins (roughly twice what was found for DZ twins). If learning subject experiencer passives is highly dependent on experience, one expects environmental factors to be the main determinant, with heritability playing a limited role. If subject experiencer comprehension, however, is governed by maturational development, one predicts that shared environment should play little role, while heritability should be very important (i.e. MZ twins should score more similarly than DZ twins). Indeed, in their analyses of shared environment versus heritability, they found a strong shared environment effect only in the case of actional passives (where good performance can be taken as due to an environmentally-determined strategy), while a strong heritability effect obtained in the subject experiencer passives, which showed no effect of shared environment. These genetic effects are exactly as predicted by maturational accounts.

Lots of recent evidence bears on the claim that children's early passives are adjectival in nature. The availability of the adjectival strategy for English-speaking children is taken to be due to the fact that the adjectival passive form in English is an s-homophone (i.e. is homophonous) to the verbal passives. The relevant prediction, then, is that in a language in which adjectival and verbal passives do not share a common pronunciation (i.e. are not homophonous), the adjectival passive is unlikely to serve as an s-homophone for the verbal passive, and therefore children speaking such a language should be equally delayed in their comprehension of both actional passives and subject experiencer passives.

Greek is just such a language. Greek adjectival passives (copula plus participle) (27) are not homophonous with Greek verbal passives (no auxiliary, and a particular verb form) [Greek verbal passives are synthetic, in the sense that they do not employ auxiliaries but are formed with N(on)Act(ive) inflection on the verb](28).

(27) To vivlio ine diavas<u>meno</u> (apo olous tous neous fitites)

the book is read (by all the new students)

(28) To vivlio diavas<u>tike</u> (apo tous fitites).

the book read-3s-NAct (by the students)

'The book was read (by the students).'

In addition to the above syntactic differences, the verbal passive and adjectival passive further differ in Greek based on agreement properties. Verbal passives are inflected for number, person, and tense. Adjectival passives are inflected for gender, case, number, and agree with the modified noun and its determiner.

Using a two-choice sentence-picture matching task, Terzi and Wexler (2002) were able to investigate young Greek children's comprehension of adjectival passives, and verbal passives with both actional and subject experiencer verbs. The same experimental materials were used with all three sentence types, and the same actional verbs were used as verbal passives and adjectival passives. They found adjectival passives to be comprehended extremely well, even in children as young as three years-old. As expected, subject experiencer passives were found to to be quite delayed, even in the five-year-olds. Unlike in English, however, actional verbal passives in Greek are quite poorly comprehended (only 44% correct in the five-year-olds group).

The extremely poor performance on the verbal actional passives demonstrates that Greek children do not use an adjectival structure, which the study shows they comprehend, in order to interpret the full verbal passive. This provides support for the hypothesis that the adjectival passive must be an s-homophone of the verbal passive for the adjectival passive to be used in that way. Where English children perform quite well on actional passives, Greek children have many difficulties. The lack of an s-homophone for the Greek verbal passive and subsequent poor performance even on actional verbal passives supports a grammatical problem for verbal passives in early child grammar.

Evidence supporting children's use of the adjectival strategy for analyzing verbal passives when such a strategy is available comes from a Russian acquisition study by Babyonyshev and Brun (2004). Russian verbs (including verbal passive participles) are inherently marked for grammatical aspect. Thus verbal passives can be formed from either a perfective (29) or imperfective (30) verb, where each has a unique syntactic structure, and where both forms allow an optional *by*-phrase.

(29) *Perfective passive*

Dom	byl	postroen	(škol'nikom)
house-NOM	was	built-PERF-PASS	pupil-INSTR

'The house was built (by the school boy)'

(30) *Imperfective passive*

Dom stroilsya (škol'nikom)

house-NOM built-IMPERF-PASS pupil-INSTR

'The house was being built (by the school boy)'

Babyonyshev and Brun examine both adult-produced and child-produced Russian speech for use of passives, noting whether the passives were perfective or imperfective. Informal adult speech was found to show a slight bias for imperfective passives, with perfective passives used only 44.2% of the time. Examining child-produced speech, no significant difference between the use of perfective and imperfective aspect in the active voice was discovered, just as no significant difference existed for the adults in the active voice. Yet, of the 212 unambiguously verbal passives produced by the children, 193 (91%) were perfective.

A similar asymmetry favoring child perfective passives was noted in an elicited imitation experiment wherein Babyonyshev and Brun attempted to elicit both perfective and imperfective passives. They found a significant difference between perfective passives (69.8% successful repeats) and imperfective passives (40.9% successful repeats).

To account for this asymmetry in Russian children's productions of passives, namely the overwhelming preference for perfective passives, which is neither present in adult use or in children's use of actives, Babyonyshev and Brun invoke a s-homophone explanation. They note that in Russian, just as in English, there are also adjectival passive participles. Russian adjectival passives (31) are homophonous with truncated perfective passives, but fail to license *by*-phrases.

(31) *Adjectival passive*

Dom byl postroen (*škol'nikom)

'The house was built (*by a school boy)'

In Russian, the passives that children produce (perfective passives) are homophonous (s-homophones) with adjectival passives, just as was found for English. The asymmetrical preference for perfective passives in child Russian can therefore be taken as evidence that Russian children analyze verbal passives as adjectival passives.[11] These data fit well with the findings from English and Greek, and together offer further support for a grammatical deficit that rules out verbal passives for young children, as well as the use of a compensatory linguistic strategy when available. What remains to be seen, however, is how to adequately encapsulate the specific grammatical deficit which rules verbal passives out in premature child grammar. As noted in 1.2.4, ACDH faces serious challenges, and as such, alternative grammatical accounts must be considered.

1.2.6 *Alternative Maturation Grammatical Accounts to ACDH*

While the evidence discussed in 1.2.5 strongly suggests that verbal passives are subject to maturational development, it is not clear how specifically to spell out what grammatical operations or structures are subject to maturation. ACDH is one account, but as discussed in 1.2.4, ACDH falsely predicts that even active sentences should be delayed once one accepts VPISH. To the extent that accounts are maintained in the syntactic

[11] How Russian children deal with the *by*-phrase (grammatical in the verbal passives, ungrammatical in the adjectival passives) suggests these structures are resultant stage adjectival passives, as these are also ungrammatical in English with a *by*-phrase.

literature where the subject is not base-generated in its final surface position, ACDH will continue to find itself on the wrong side of correctly predicting children's acquisition of active voice sentences (and adjectival passives under new treatments that suggest they too involve an *A*-chain; see Kratzer, 2000, Anagnostopoulou, 2003, Embick, 2004). What is therefore needed is a grammatical (maturation) account that correctly rules out verbal passives, but not actives and adjectival passives. Several different theories claiming to do just that have been posited in recent years and are reviewed below. All, like ACDH, attribute comprehension difficulties with subject experiencer passives to a grammatical deficit (due to biological/neurological genetically-determined immaturity), and success with contemporaneous actional passives to a compensatory heuristic (i.e. the adjectival strategy). What concerns us here are differences in how the grammatical deficit is formulated.

1.2.6.1 External Argument Requirement Hypothesis

An alternative proposal attempting to account for delays in verbal passive acquisition, but which allows VP-internal subjects, is Babyonyshev, Ganger, Pesetsky and Wexler's (2001) External Argument Requirement Hypothesis (EARH). According to EARH, premature child grammar does not license any structures that do not have an external argument. That is, children demand that external arguments be part of clausal projections. The authors formalize EARH in Minimalist terms under the hypothesis that young children consider structures with v_{def} to be ungrammatical. Here v is taken to be def(ective) when it does not assign an external argument theta-role. EARH posits verbal passives to involve just such a defective v on the assumption that subject position in a passive must be non-thematic. Since verbal passives are taken to have this v_{def}, they are ruled out by EARH.[12] On the other hand, EARH predicts that VP-internal subjects (or subjects in [Spec, *v*P])can raise to [Spec, IP] with no problem, since these structures do have an external argument.

1.2.6.2 Canonical Alignment Hypothesis

Hyams, Ntelitheos, and Manorohanta (2006) seek another explanation for why passives are delayed in acquisition. They write, "[d]escriptively speaking, children's difficulty seems to be restricted to those *A*-chains that derive a misalignment of thematic and grammatical hierarchies. The argument structure associated with transitive…and unergative… verbs…specifies an actor-like external argument, which is not represented in the passive, which involves a promotion of the theme to the external argument position." In essence, they believe the problem with verbal passives is that a "canonical alignment" is not respected (theme in subject position), which premature children find ungrammatical. CAH claims that children are unable to represent structures that derive a mismatch between a syntactic position and the canonical theta-role associated with that position. As such, subject position is uniquely reserved for agents, object position for themes. While adult grammar is flexible in what theta-roles are associated with landing sites, CAH hypothesizes that children's grammar is much more restrictive. Active sentences, which do respect canonicity, are predicted to pose no issue for young children on CAH.

---

[12] Collins (2005b) offers a compelling argument that the *v* in verbal passives does assign a thematic-role to the external argument, and this is how the *by*-phrase comes to have the correct theta-role. See Figure 1.2.2. If Collins is correct, then EARH incorrectly predicts verbal passives to be grammatical for young children.

1.2.6.3 Universal Freezing Hypothesis

Hyams and Snyder (2005) hypothesize that premature child grammar requires a strong version of the Freezing Principle developed in Wexler and Culicover (1980), and reformulated by Müller (1998). The Freezing Principle basically rules out movement from an already moved phrase. In putting forward their Universal Freezing Hypothesis (UFH), Hyams and Snyder also adopt Collins' (2005b) analysis of verbal passives that makes use of "Smuggling". Collins attempts to account for how in certain cases a DP appears to *A*-move past another DP, which would at first seem to be a violation of the Minimal Link Condition (MLC; Chomsky, 1995) or of Relativized Minimality (RM; Rizzi, 1990).

Under Collins' (2005b) analysis of passives, the subject is generated in a *by*-phrase in canonical external argument position ([Spec, *v*P]). The object must therefore raise above the (subject-containing) *by*-phrase if it is to reach subject position, without violating the MLC or RM. This, Collins theorizes, is accomplished by means of Smuggling. Under his analysis, the VP containing the object first raises above the external argument, whose movement does not violate either the MLC or RM, after which the object is free to move to subject position. The object is thus "smuggled" past the subject, before itself moving higher (Figure 1.2.2).[13] Smuggling of this sort is an exception to the Freezing Principle since the object is raising from a VP that itself has already raised.

[13] Collins' (2005b) analysis of verbal passives does require Smuggling even in cases of truncated passives, as he hypothesizes that the object DP must still raise past a PRO external argument.

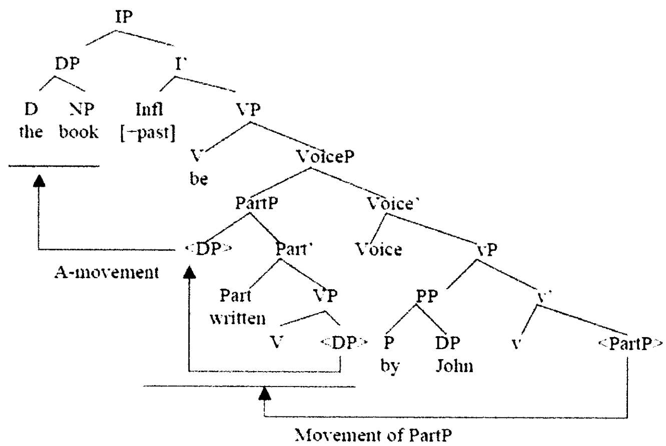


Figure 1.2.2: Phrase structure diagram of a verbal passive involving Smuggling (copied from Collins, 2005b).

1.2.6.4 Universal Phase Requirement

Wexler (2004) proposes the Universal Phase Requirement (UPR) as a replacement for ACDH. UPR is couched in the Minimalist Theory framework and relies crucially on certain assumptions about adult grammar found in Chomsky (1998, 2001). Chomsky derives on Minimalist considerations a very strong cyclic theory of syntax. Essentially Merge proceeds from the bottom to the top of a derivational tree with most of the derivation closed off to further analysis or change as it proceeds. Chomsky proposes the Phase Impenetrability Condition (PIC) in (32):

(32) PIC: When working at a phase, the edge (the head and any specs) of the next lower phase is available for analysis, but nothing lower than the edge. In particular the complement is not available.

C (Complementizer) and *v* (the light-verb head of a phrase that takes VP as its complement and selects the external argument as its specifier) are substantively defined as phasal heads (on an argument from completeness of semantic properties).[14] Subjects of *v*P can move to [Spec, IP] because the subject, although inside a phase *v*P, can be probed at the next phase CP because the subject is at the edge (specifier position) of *v*. Passives, however, create a special problem for the theory, as the object of V must move up to T (INFL) in the higher phase. Chomsky proposed that the *v* of passives is "defective," that is, it does not assign an external argument and it does not operate as a phase. Thus the object of V is available at the higher C above it. The *v* of transitives (including unergatives) is denoted v^*; it is this *v* that both selects an external argument in its specifier position, and defines a phase.

UPR holds there is no head v_{def} in premature child grammar, that all instances of *v* universally define a (strong) phase for children. UPR states that children take the *v* of passives to be v^*_{def}. That is, for premature children, the *v* in passives (i.e. v^*_{def}) is phasal (hence "*"), but does not assign an external argument (hence "def").[15] As such, verbal passives are ungrammatical for the premature child because the object cannot move past the strong phase boundary v^*_{def}. This is the correct result empirically (ignoring impersonal passives for the sake of generality; see Verrips, 1996). On the other hand, the "VP-internal subject" of a transitive (i.e. one with an external argument) clause,

[14] This and other aspects of syntactic theory that are crucial to UPR are taken up in Section 4.1.3.

[15] To be clear, UPR makes the claim that children have access to a head (i.e. v^*_{def}) that is not part of adult grammar. UPR cannot claim that children simply treat all instances of *v* as v^*, as this would rule out structures involving (Merged) expletives (e.g. unraised expletive sentences, weather-*it*, etc, which are known to be early in acquisition; see Section 2.2.4). An alternative analysis that disconnects the phasal and theta-role/argument assigning aspects of *v* (for motivation, see Harley, 2007 and Schäfer, 2008), might be able to capture the same predictions without relying upon the need for an otherwise unmotivated head as v^*_{def}. For the purposes of further discussion in this work, I will accept UPR as originally put forth in Wexler (2004).

generated in [Spec, *v*P] is at the edge of VP. At the next higher phase, C, the subject is available for analysis according to PIC. Thus T can "see" the subject, meaning that Agree and Move can take place; there is no need for non-phasal *v*. The child subject to UPR is also unhindered since there is no non-phasal *v* involved. UPR predicts that the child can raise the subject from the edge of *v*P with no problem. This solves the problem of VP-internal subjects that dogged ACDH.

The upshot of UPR is that there is no problem with any particular kind of chain, and no special assumptions about objects or subjects moving, or even Agreeing. The problematic constructions are those that demand movement across a defective phase (e.g. non-phasal *v*) in the adult grammar.

These four maturation grammatical theories (EARH, CAH, UFH, and UPR) all correctly predict children's delayed comprehension of verbal passive. What is needed in order to validate one over the others (or to tell us if a new theory altogether is needed) are other grammatical structures for which they make differing predictions. By examining how each fares in predicting the timeline of acquisition for other grammatical structures under their purview, one will better be able to sort out which theories are worth holding on to for further consideration, and which merit abandonment. One very attractive grammatical case that could help differentiate between these four grammatical acquisition theories, and that until now has gone unexamined in the acquisition literature, is that of subject-to-subject raising, to which we now turn.

1.3 The Case for Subject-to-Subject Raising

In order to help validate the grammatical acquisition accounts for verbal passive delay discussed in Section 1.2.6, new data are needed. In particular, what must be sought is evidence concerning the acquisition of grammatical operations/structures for which those grammatical accounts make differing predictions. One such syntactic phenomenon, which has hitherto received extremely limited investigation in the field of first language acquisition is that of subject-to-subject raising.

1.3.1 *Subject-to-Subject Raising in Syntactic Theory*

Subject-to-subject (StS) raising refers to the grammatical phenomenon in which a matrix subject serves as the logical (and structural) subject of an embedded clause, while having no semantic/thematic relationship with the matrix (StS raising) verb (33).

(33) John seems to walk in the garden.

Thus, in the above sentence, the matrix subject *John* is understood as the person that walks in the garden (*John* is a "walker"), even though *John* appears immediately to the left of the verb *seem*, from which it receives no thematic-role (*John* is not a "seemer"). StS raising verbs, thus, are one-place predicates that take a clausal complement and assign no θ-role. Examples of StS raising verbs include: *seem*, *appear*, *tend*, *used*, *begin*, *need*, and *continue*.

A sentence like (33) can be differentiated from a subject control sentence like (34), where both share the same surface structure in English (subject + verb + embedded nonfinite clause), but have different syntactic and semantic realizations.

(34) John wants to walk in the garden.

In this sentence, the matrix subject *John* is taken to be the logical subject of both *want* and *walk* (*John* is both a "wanter" and a "walker"), even though it appears phonologically only to the immediate left of *want*. In the case of subject control (34), *John* is a semantic argument of both the matrix and embedded verbs, while with StS raising (33), it is a semantic argument only of the embedded verb. Examples of subject control verbs include: *want*, *try*, *ask*, *prefer*, *remember*, *expect*, and *hope*.

To help further clarify the difference between (33) and (34), consider the following sets of sentences. First, if the infinitival clause allows an expletive/dummy subject (35a), then the raising verb will allow an expletive subject (35b), while the control verb will not (35c).[16]

(35) Expletive Subjects (Expletive-Associate Construction)

a. There is a man in the garden.

b. There seems/appears/tends/used/begins/needs/continues to be a man in the garden.

[16] While all StS raising verbs allow for an expletive subject with an infinitival embedded clause, the ability to have a dummy subject with finite embedded clauses is limited to only some StS raising verbs:

(i) It seems/appears/*tends/*used/*begins/*needs/*continues that John walks in the garden.

c. *There wants/tries/asks/prefers/remembers/expects/hopes to be a man in the garden.

Second, consider an idiom like (36a), where the meaning of this sentence is that a secret has become known. The idiomatic interpretation is maintained in the raising sentence (36b), but lost in the subject control sentence (36c), where only a literal interpretation is possible.

(36) Idiom-chunk Subjects

a. The cat was let out of the bag.

b. The cat seems/appears/tends/used/begins/needs/continues to have been let out of the bag.

c. *The cat wants/tries/asks/prefers/remembers/expects/hopes to have been let out of the bag.

Third, one can ask if truth-conditional equivalence is maintained under passivization. Just as the active and passive monoclausal sentences in (37a) have the same truth-conditions, the raising sentences in (37b) also have the same truth-conditions regardless of whether the embedded infinitival is in the active or passive voice. This differs greatly from (37c) where the two subject control sentences are not truth-conditionally equivalent when the voice of the embedded clause is switched (i.e. John wanting to love Mary is certainly not the same as Mary wanting John to love her).

(37) Truth-conditional Equivalence with Voice Alternations

a. John loves Mary = Mary is loved by John

b. John seems to love Mary = Mary seems to be loved by John

c. John wants to love Mary ≠ Mary wants to be loved by John

What these various diagnostics demonstrate is that the matrix subject in the raising sentences receives all of its meaning from and relative to the embedded clause, and not from the raising verb. In the case of the expletive subjects, with a matrix raising verb the embedded clause subject need not even appear in the matrix clause since the subject is fulfilling its syntactic and semantic roles *in situ*. With idioms, the embedded clause subject retains its idiomatic interpretation in the raising sentences since no semantic relationship obtains between it and the matrix raising verb. Finally, the reason that active and passive embedded clause are semantically the same under raising verbs is that the raising verb does not semantically relate to the matrix subject.

These various factors as relates to StS raising have traditionally been captured in the (Government & Binding) syntactic literature by positing that StS raising verbs take a single clausal argument and assign no theta-role.[17] The subject position of a raising verb is left unspecified (38).

(38) $[_{IP}$ e $[_{VP}[_{V'}$ seems $[_{IP}\ [_{NP}$ John$]\ [_{VP}\ [_{V'}$ loves $[_{NP}$ Mary$]]]]]]]$

Given a syntactic principle requiring all sentences to have a (not necessarily phonologically overt) subject (e.g. EPP), in (38), the matrix subject position may either

[17] For a brief discussion on how subject control is captured syntactically in GB, see Section 3.3.1, or for a more general review, see Haegeman (1994). Section 4.1.2 briefly discusses control in Minimalist Theory (mostly as relates to the theory of Control as Raising).

be filled by merging an expletive (i.e. *It*) as in (39) or by raising a lower argument (i.e. *John*), which leaves behind a (co-indexed) trace as in (40).

(39) It seems (that) [John loves Mary]

(40) John_i seems [t_i to love Mary]

Note that it is only the subject of the embedded clause that can raise to the matrix subject position with a matrix StS raising verb (40), and not another argument from the embedded clause, such as a direct object (41).

(41) *Mary_i seems [John to love t_i]

This ability of the subject to *raise* from the subject position of the embedded clause to the subject position in the matrix clause is what gives this syntactic phenomenon the name *subject-to-subject raising*. It is the acquisition of this particular type of syntactic movement that is the focus of this dissertation.

The "traditional" syntactic account in Government and Binding for this sort of movement has the following logic (e.g. see Haegeman, 1994 for a general review). Since the StS raising verb does not specify an external argument, the matrix subject position is left empty (see (38) above). Given a syntactic requirement that *something* must come to fill this position two options present themselves. Merging an expletive is one (39), and the other option is found in (40), where the closest element of the appropriate type raises to the matrix subject position (where "closest" and "appropriate type" are in need of formal definition).[18] This movement of an argument (the embedded subject) to another argument position (as the matrix subject) is a form of *A*-movement, as was seen for verbal passives in Section 1.2.1.

One additional difference between StS raising verbs and subject control verbs that will become relevant in the acquisition discussions to follow is that some verbs in the former verb class can take an optional experiencer-phrase (42), while the latter cannot (43).

(42) John seems/appears <u>to me</u> to dance (every Saturday).[19]

(43) *John wants/tries/asks/prefers <u>to me</u> to dance (every Saturday).

The experiencer-phrase that appears with StS raising verbs in English is (usually) headed by the preposition *to*.[20] Further details concerning the syntactic nature of experiencers in StS raising structures are taken up in Section 3.1.2.

1.3.2 *What the Study of Raising Can Offer*

An experimental investigation into children's acquisition of StS raising offers clues to several important research questions in the fields of language acquisition (and

---

[18] It is not locality alone that determines which element raises. Consider (38). Even if the embedded subject and object were of equal "closeness", the object would still not raise. The object must check case within the domain of the embedded infinitival, while the subject cannot check case in that domain, so it alone must raise to the higher subject position.

[19] Many other StS raising verbs, however, do not license an experiencer-phrase (i):

(i) #John tends/used/begins/needs/continues <u>to me</u> to dance (every Saturday)

[20] The StS raising verb *strike* can also take an experiencer-phrase, but it is not headed by *to* (i) and (ii):

(i) John strikes me to be a fine fellow.

(ii) It strikes me that John is a fine fellow.

linguistics more generally), as well as psychology, neuroscience, and biology. As relates to first language acquisition in typically developing children, evidence concerning the time course of development of StS raising is conspicuously absent, *especially* given that this syntactic phenomenon is a rather central focus in the generative syntax literature. As an example, Robert Borsley's syntax textbook (the first introductory text I grabbed from my shelf) *Syntactic Theory* (1999) has 13 chapters dedicated to various important aspects of syntax, covering such topics, to name just a few, as heads and complements, anaphora, passives, control, *wh*-dependencies, and island constraints. Given 13 topics that Borsley found important enough to merit an entire chapter, the only one that does not also receive coverage in Theresa Guasti's excellent and comprehensive child acquisition textbook *Language Acquisition: The Growth of Grammar* (2002) is that of raising. That such an important research area in adult linguistics has yet to be investigated in child acquisition is clearly an oversight in need of remedy.

As relates to learning theories in psychology and the cognitive sciences, this research also serves as a further test of the viability of maturation theories as explanations for language acquisition delay. If such accounts are to hold for StS raising, as was the case with the maturation theories posited to account for delays in children's comprehension of verbal passives, they entail that similar cross-linguistic, environmental, and genetic evidence should eventually be found as corroboration. More specifically, StS raising might help distinguish between various maturation accounts that correctly predict delayed acquisition of verbal passives, but for different syntactic reasons.

Let us consider the grammatical acquisition theories discussed in Section 1.2.6. They attempt to capture the linguistic difference(s) between premature child grammar and adult grammar that renders verbal passives ungrammatical for premature children and ask what these accounts predict for StS raising. First, there is EARH. On this account, structures containing clauses that do not project external arguments (i.e. those with v_{def}) are taken to be ungrammatical for premature children. Since both unraised and raised sentences with StS verbs contain a v_{def}, EARH predicts that children should find both ungrammatical. That is, according to EARH, young children should not license sentences like either (39) or (40) above.

On CAH, those structures that have arguments which appear in positions with atypical theta-roles are ruled ungrammatical. Verbal passives (with a theme in subject position) are an example of one such violation. As applies to an unraised sentence like (39), CAH predicts no difficulties for premature children. The expletive subject receives no theta-role at all, so cannot derive a mismatch between position and theta-role. To the extent, then, that the embedded clause does not violate CAH, unraised sentences should be fine. Likewise, CAH does not obviously predict any difficulties for raised sentences like (40) either. In particular, if the embedded clause assigns its subject an agent theta-role, then when the subject raises to the matrix subject position, it will in no way challenge any mapping rules.

UFH predicts problems only for those structures in adult grammar than involve movement from already-moved phrases (i.e. those that are exceptions to the Freezing Principle). As relates to verbal passives, their delay in acquisition is captured on UFH, since by Collins' (2005b) account, they involve Smuggling (to get the direct object past the c-commanding *by*-phrase). The grammaticality of unraised and raised sentences will therefore depend on whether or not they, too, involve *A*-movement of one argument past another (requiring Smuggling). Unraised sentences like (39) do not involve any such syntactic operations, and are predicted to be just fine for young children on UFH. Turning to raised sentences, the issue is a bit more complicated, since according to Collins (2005a), the syntactic analysis depends on whether or not the raised sentence contains an experiencer-phrase. Consider a sentence like (42). According to Collins, its derivation involves Smuggling the embedded subject past the experiencer. Consider the derivation for the sentence in (44).

(44) John seems to Mary to be nice.

According to Collins (2005a), the derivation proceeds as follows. First, the infinitival embedded clause is built, resulting in (45). Then the StS raising verb *seem* is introduced, with the embedded subject then moving to [Spec, VP] (46).[21] Next, there is extraposition of the infinitive (47). At this point the experiencer is introduced, assumed by Collins to be in the specifier position of an Appl(icative) phrase, and then the light verb *v* is merged (48). Now the crucial step occurs, whereby the VP moves clause-internally to [Spec,*v*P] (49).[22] This Smuggles the embedded subject (inside the VP) over the experiencer. Finally, the matrix IP is built and the subject raises to matrix [Spec, IP] (50). The full derivation is found in Figure 1.3.1.

(45) $[_{IP}$ John to be nice]

(46) $[_{VP}$ John $[_{V'}$ seem $[_{IP}$ <John> to be nice]]]

[21] This step is a bit suspect (see 4.1.3). It involves movement without any Agree relation obtaining between *seem* and *John*.
[22] This movement is taken to be driven by an uninterpretable [uV] feature and something like EPP.

(47) $[_{XP}\ [_{IP}\ \langle \text{John} \rangle \text{ to be nice}]\ [_{X'}\ X\ [_{VP}\ \text{John}\ [_{V'}\ \text{seem}\ \langle IP \rangle]]]]$

(48) $v\ [_{ApplP}\ \text{Mary}\ [_{Appl'}\ \text{Appl}\ [_{XP}\ IP\ [_{X'}\ X\ VP]]]]$

(49) $[_{vP}\ VP\ [_{V'}\ v\ [_{ApplP}\ \text{Mary}\ [_{Appl'}\ \text{Appl}\ [_{XP}\ IP\ [_{X'}\ X\ \langle VP \rangle]]]]]]$

(50) $[_{IP}\ J\ [_{I'}\ I\ [_{vP}\ [_{VP}\ \langle J \rangle\ [_{V'}\ \text{seem}\ \langle IP \rangle]]\ [_{v'}\ v\ [_{ApplP}\ M\ [_{Appl'}\ \text{Appl}\ [_{XP}\ IP\ [_{X'}\ X\ \langle VP \rangle]]]]]]]]$

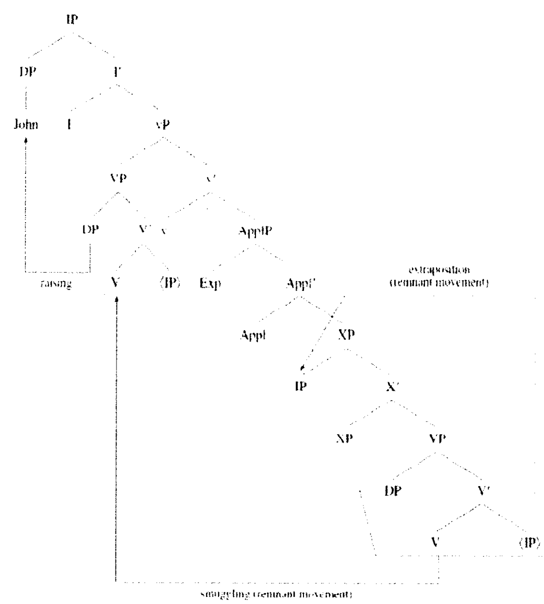


Figure 1.3.1: Phrase structure diagram of StS raising over an experiencer-phrase involving Smuggling (copied from Collins, 2005a).

Importantly, Collins argues that the above derivation, with remnant movement, extraposition, and Smuggling, only occurs when an experiencer is present (since those operations are only required to account for how the embedded subject comes to move past the experiencer). In a raised sentence like (40), where there is no experiencer-phrase, there is no Smuggling. On UFH, then, raising should be possible when there is no experiencer, but is ruled out when an experiencer is present.

On UPR, structures with defective phases *can* be unproblematic for premature children. Consider an unraised structure like (39). For the child, the matrix v must not assign an argument (i.e. there is an expletive subject), but due to UPR, must be phasal: this is v^*_{def}. There is no reason, however, that the child cannot make the derivation converge. No relation holds between matrix T and anything in the lower clause, so that even if the child takes v to be phasal, nothing in the computation is interrupted. UPR predicts that unraised *seem* in sentences should converge for the immature child, even if the derivation used (with phasal v^*_{def}) is slightly different than the adult representation.

On the other hand, consider Chomsky's (2001) analysis of the sentence in (44), as is seen in (51).

(51) John T v_{def} seems to Mary [t_i T_{def} to be t_i v^* nice]

In adult grammar this analysis requires that matrix v be non-phasal (i.e defective) in order for the embedded subject to be accessible for movement to the higher clause, since it does not sit at the edge of the lower clause. By UPR, the child takes matrix v in (51) to be phasal. T then cannot probe *John* (which is in the complement of v), and T will end up with uninterpretable features unchecked. As such, the raised sentence should be ungrammatical for young children. The presence or absence of the experiencer-phrase should be irrelevant on UPR, as in either case, movement should be ruled out if the raising v is phasal.[23]

[23] Note that for UPR to make this prediction, Collins (2005a) cannot be correct. That is, if Smuggling does take place when raising over an experiencer, the embedded subject does first move to the edge position before raising to the higher clause, and as such, should not be ruled out (raising with no experiencer-phrase would still be problematic). See 4.1.3 for discussion.

1.3.3 *Agenda*

Chapter 1 has offered a brief survey of current understanding regarding the acquisition of verbal passives and laid out the case for why studying the acquisition of StS raising is of great importance, as applies both to theories in language acquisition and fields beyond. Chapter 2 investigates StS raising with an experiencer-phrase, and how the acquisition of such structures relates to the acquisition of verbal passives. Chapter 3 addresses potential limitations and concerns in studying StS raising with an experiencer-phrase, and sets out to investigate StS raising with no experiencer phrase, as well as how acquisition of such structures relates to verbal passives. Finally, Chapter 4 considers production data of StS raising structures (both child-directed and child-produced), additional grammatical structures for which relevant grammatical acquisition theories make predictions, and relevant developments in linguistic theory that bear on the adequacy of these various acquisition accounts. The chapter ends with a brief discussion of how the newly unearthed data fit into broader theories of language acquisition, psychology, neuroscience, and biology. The work's conclusions are that StS raising (both with and without an experiencer-phrase) appears delayed in typically-developing children, that the time course of development of StS raising appears to mirror that of verbal passives, and that these data are best captured by a grammatical acquisition account positing that delays are due to children's inability to manage movement across phase boundaries.

Chapter 2: Comprehension of Raising (over an Experiencer)[1]

2.1 Comprehension of Subject-to-Subject Raising

When embarking on an investigation into generally uncharted scientific domains, it is prudent to have the end goal at the forefront of one's mind, and then proceed to take carefully planned baby steps. At the start of this inquiry, little existed in the acquisition literature to help guide discovery and unearth answers to the question of what children know concerning syntactic subject-to-subject (StS) raising,. As discussed in Chapter 1, a guiding reason to undertake an investigation into children's comprehension of raising is to get a better sense of why young children fail to comprehend certain syntactic phenomena (e.g. verbal passives), but not others (e.g. *wh*-movement). While it certainly would be nice to know when children acquire all sorts of grammatical structures, the study of raising as explored here was not undertaken simply to add a data point to the proverbial (and totally fictitious) "grand chart of when children acquire various grammatical operations". This research has been undertaken in order to validate, reject, and improve upon theories of acquisition that already exist and already make predictions concerning children's acquisition of raising. We come to acquisition questions of raising *through* pre-existing theoretical work, with the hope of being able to amend and refine those theories, and if need be, offer wholly original accounts if new data is incompatible with previous hypotheses.

[1] Much of this chapter draws upon research published in Hirsch and Wexler (2007a). In particular, Studies 1 and 3 appear in that work. New individual subject analyses have been included for Study 1, while discussion has been greatly expanded for both studies. Study 2 is presented in print for the first time. All experiments were conducted after careful discussion and planning with Ken Wexler, who also contributed heavily to the interpretation and conclusions.

In designing a first experiment to investigate children's knowledge of StS raising, fundamental preliminary issues fell into three main areas: finding a means to differentiate syntactic knowledge of StS raising from lexical knowledge of StS raising verbs, determining whether children comprehend unraised forms of StS verbs and how unraised sentences could be used as controls for investigating acquisition of StS raising, and lastly deciding what aged children should be studied.

2.1.1 *Knowledge of Raising vs. Knowledge of Raising Verbs*

A challenge facing researchers investigating children's knowledge of StS raising is how to distinguish a failure to comprehend raising from merely a failure to recognize a particular verb as a StS raising one. For example, if children failed to comprehend the syntax and semantics of a sentence like (1), would one want to conclude on that basis that children lack grammatical StS raising?

(1) John seems to wash his car every Saturday.

Could it not be the case that the child simply has never heard this particular verb (i.e. *seem*) before, and thus does not even bother to attempt a grammatical parse? Or that he has knowledge of raising, but for whatever reasons, has miscategorized this raising verb as a control verb? This latter possibility needs be seriously considered given that (in English at least), the vast majority of verbs fitting the frame *X VERB to Y* are control verbs, not raising verbs.[2] Furthermore, while some raising verbs fitting this form are

[2] While no detailed counts were made to verify this claim, a quick back-of-the-napkin listing easily produces more control than raising verbs fitting this pattern.

relatively frequent in the input to children (e.g. *seem*; see Chapter 4), others are much less so, while many of the subject control verbs are extremely frequent (e.g. *love*, *hate*, *want,* etc). Thus the question of when children acquire raising is entangled with the related, but different question of how children learn to correctly categorize particular verbs as raising or control verbs. The latter forms the basis of Becker's (2005, 2006) research, to which we will turn in Chapter 3. In the meantime, any initial experiments one conducts to investigate the former question (i.e. when is StS raising acquired) must include some means of verifying how children categorize these verbs independently, to whatever extent possible, of the grammatical operations that can be applied to structures involving them.

Considering how children could respond to raising sentences can lead to better-designed comprehension experiments. If children undergoing testing do have adult knowledge of grammatical raising, what could lead them to misanalyze a particular raising sentence? First, the child might not know the particular raising verb being used in the test sentence. Perhaps he has never heard it before, or has not heard it frequently enough to define it. In this case, the child might do nothing better than guess at the meaning of the sentence, perhaps resulting in random chance behavior during an experiment. Second, the child might have misdefined the verb, so that he thinks he knows its meaning, but happens to be wrong. As discussed above, one potentially likely miscategorization would be to treat the raising verb as a control verb. With some insight into what semantics the child applied to this 'raising as control' verb, perhaps experimental manipulations could be designed to catch just such mistakes.

Control: *want, remember, prefer, fight, work, hope, pray, claim, dream, wish, plead, slave, hate, love, rally, beg, promise* (with its own interesting orthogonal issues, see Larson, 1991), *care, expect, hope, detest, loathe, try, attempt, yearn*, among many, many others.
Raising: *seem, appear, tend, happen, used, begin, need, continue.*

One must also consider how children would behave in case they did lack StS raising. Having no ability to implement the required syntactic operations, children might simply give up when attempting to parse a sentence involving raising. This might lead to chance performance given a particular experimental paradigm. Children might instead attempt to parse the sentence assuming the verb is of the non-raising variety. Once again, children might assume raising verbs are control verbs, which share similar syntactic frames with the raising verbs in many contexts. Here too, one could attempt to distinguish such parsing patterns given certain experimental manipulations. A great challenge, to which we will return, would involve determining if and why some children treat StS raising verbs as something else, distinguishing between the possibility that they know raising, but have miscategorized raising verbs as other kinds of verbs, and, alternatively, that they lack raising, and therefore have miscategorized raising verbs as other verbs.

Given the potential concern about differentiating possible failure due to lack of raising knowledge versus failure due to verb miscategorization, there are a few clear paths researchers can take in designing initial studies to probe the acquisition of raising. First, researchers should choose StS raising verbs that give children the greatest possible chance of obtaining adult parses. This would involve selecting only the most frequent raising verbs in the input to children; ideally, verbs that children demonstrate that they use at an early age. Given the discussion concerning adult and child productions in Chapter 4, it is clear that in the case of the raising verb *seem*, children both hear and produce it from very early ages (whether child productions involving *seem* are evidence for grammatical raising is taken up in that chapter). The raising experiments in Chapter 2 will therefore limit themselves to this single raising verb, which given its frequency in adult-to-child input should positively bias the chance of children successfully interpreting raising sentences containing it. Using a single raising verb in these first experiments also adds a certain level of simplicity and control to the experimental designs. The choice of a single verb serves to simplify the experiment by minimizing extraneous factors (e.g. lexical learning, verb differentiation, etc).

Using only high-frequency raising verbs, and even limiting oneself to only a single high-frequency raising verb, however, does not guarantee that any individual child would know the particular raising verb chosen, in this case *seem*, regardless of its general input frequency across many children. A researcher would therefore greatly appreciate a secondary means of inferring whether children know something about a particular raising verb prior to testing a child's knowledge of that verb in sentences involving raising. One means of inferring whether children recognize a verb to be a StS raising verb without necessarily having acquired the syntactic principles governing StS raising is to test children's knowledge of the raising verb in its unraised form. As discussed in Section 1.3.2, grammaticality involving raising verbs can be satisfied in one of two ways: either through Merge (of an expletive in the specifier position of the raising verb) or Movement (of the embedded subject DP to the specifier position of the raising verb). Only the latter involves the syntactic raising at the heart of this dissertation. If one can show that children comprehend a raising verb in its unraised/Merged variant (2), but not its raised/Moved variant (3), then one can reasonably conclude that problems with raising are not tied to lack of lexical knowledge of the raising verb, but rather that grammatical raising itself is delayed.

(2) It seems/appears (that) Jennifer is dancing.

(3) Jennifer seems/appears to be dancing.

2.1.2 *Knowledge of Raising Verbs in their Unraised Forms*

As just mentioned, the inclusion of unraised sentences involving StS raising verbs provides a control for interpreting child failures on raised sentences involving the same verbs. If a child fails to comprehend both raised and unraised sentences involving a particular raising verb, one would not want to conclude that the child necessarily lacks syntactic StS raising. Two alternative possibilities seem clear. First, the child might actually have already acquired the grammatical principles underlying knowledge of syntactic raising, but nonetheless not have learned the raising verb being tested. As discussed in 2.1.1, the child might not have heard the verb used sufficiently to define it correctly as a raising verb (e.g. instead treating it as a control verb), or perhaps not having heard it enough to attempt any definition at all. In either case, failure would thus be predicted and expected on both unraised and raised forms involving this verb. Since in this case we are hypothesizing that the child has already acquired raising, failure on the unraised form is attributed to an issue with lexical acquisition. Second, the child really could lack syntactic raising, thus straightforwardly accounting for failure with the raising form, but then, what about failure on the unraised form? This could again be due to lack of lexical knowledge (not recognizing the verb as a raising verb, independent of knowledge of raising), or to a grammatical hypothesis on the part of the child that since raising is not grammatical, raising verbs cannot exist (regardless of whether they are realized in a raised or unraised structure). Such children lacking raising might hypothesize that all verbs fitting the frame *X VERB to Y* must be theta-assigning control verbs. In this case, both the unraised and raised forms would be ruled ungrammatical: the unraised forms since they fit no known grammatical structure (expletive structures are non-theta-assigning), and the raised forms since the matrix subject receives its theta-role from the embedded clause only.[3] Such children might rule raised sentences to be ungrammatical (i.e. disallowed by grammar) or simply assume they are control structures; children's actual analysis would have to be empirically determined.

Using unraised forms as a sort of lexical control for interpreting children's knowledge of raising sentences is predicated upon children finding unraised sentences grammatical.[4] Whether children do in fact consider unraised forms grammatical has yet to be subjected to empirical verification in the acquisition literature.[5] In fact, the various maturational grammatical theories discussed in Chapter 1 concerning passive acquisition make different predictions with respect to the timing of acquisition for unraised sentences. Most of these theories predict children should find unraised sentences grammatical at ages when raising is predicted to be ungrammatical. That is, they predict earlier acquisition for unraised sentence forms compared to raised sentence forms. The Canonical Alignment Hypothesis (CAH; Hyams, Ntelitheos, and Manorohanta, 2006),

[3] This ignores (for the moment; see later sections of this chapter for discussion of this issue) the possibility that a child having not yet acquired raising might interpret a (StS) raising verb as either a Raising to Object (RtO) verb or a control verb, in which case the (StS) raising sentence form itself would be grammatical, but would involve a non-(StS) raising (i.e. RtO or control) syntactic analysis and interpretation.

[4] This is itself complicated experimentally by the fact that even if children find unraised forms grammatical, there will be some who could nonetheless fail on comprehension experiments testing such forms due, again, to independent reasons (e.g. failure to have acquired the particular raising verb). As always, researchers must judiciously interpret children's incorrect responses to particular experimental stimuli. A failure to comprehend specific test stimuli designed to reflect certain grammatical phenomena need not signal a failure to have acquired those phenomena. Sometimes a cigar is just a cigar.

[5] This is true with regards to published research. The unpublished study by Froud, Tsakali, and Wexler described briefly in Section 2.1.4 did indeed examine children's comprehension of such unraised sentences, generally finding children performed quite well on them. In particular, these researchers found children performed much better given unraised sentences compared to their semantically equivalent raised versions. This empirical dichotomy is support for Wexler's (2004) UPR.

also discussed in Chapter 1 in relation to children's early comprehension of verbal passives, predicts no problems for unraised forms. This is because the only relevant external argument, namely the expletive Merged into matrix Spec,TP, bears no theta-role at all, and thus is not subject to redefinition of the linking rules governing the mapping of theta-roles onto various argument positions. Hyams and Snyder's (2005) Universal Freezing Hypothesis (UFH) would allow unraised sentences since they do not involve the movement of an already moved phrase, and thus Freezing does not apply. The Universal Phase Requirement (UPR; Wexler, 2004) has immature children treating all instances of *v* as defining a strong phase, thus ruling out any weak phases for such children. Unraised sentences for children subject to UPR would therefore have a non-adult analysis, namely involving v^*_{def}, but would nonetheless be grammatical, since the matrix expletive is Merged, not Moved past the hypothesized strong *v* phase boundary.

The final grammatical acquisition account of passives discussed in Chapter 1 that makes predictions for the acquisition of StS raising is the External Argument Requirement Hypothesis (EARH; Babyonyshev, Ganger, Pesetsky, and Wexler, 2001). Unlike the other theories, however, EARH predicts unraised forms to be ungrammatical for premature children. As reviewed in Section 1.3.2, StS raising verbs, regardless of whether realized in unraised or raised sentences, involve defective *v*, that is, a light verb that does not select an external argument. EARH is defined by the hypothesis that children consider structures involving defective *v* to be ungrammatical. As such, EARH not only predicts verbal passive and raising structures to be ungrammatical for premature children, but rules unraised forms to be likewise ungrammatical, for the same reason, namely, the presence of defective *v*.

Initial experimentation attempting to address early knowledge of StS raising, therefore, should not just include raising verbs in their raised form, but should also test such verbs in their unraised counterpart. The reasons for doing so are twofold. First, as discussed in Section 2.1.1, the use of the unraised sentences provides a lexical control for interpreting failure of raised sentences. Since this holds, however, only to the extent that unraised structures are grammatical for young children, it brings us to a second reason for including unraised forms. As discussed in this section, it is an empirical question whether or not the unraised variants are indeed grammatical for young children. Before one can make use of the unraised sentences as controls for the raised sentences, it must be experimentally verified that children do in fact find such sentences to be allowed by their linguistic system. The ultimate reason, though, for also testing unraised sentences is the same as that underlying the testing of raised sentences in the first place: determining the correct theoretical characterization of children's early grammar, and charting the developmental course of acquisition for that grammar.

2.1.3 *Relevant Children*

Choosing apposite participants for an experimental study is often just as important as the design of the experiment itself. First, one must find participants both willing and capable of completing the tasks set before them. Second, in the case of acquisition studies, children of the appropriate ages must be tested. They must be old enough to have the general cognitive capacity to comprehend the task demands of the particular study. In the case of comprehension studies, this often means being able to sit patiently while listening and watching moderately complex scenarios, which regularly involve following a narrative, watching puppets play out a scene, or attending to different pictures, after which the children are required to make some decision about what they have just seen. During this time, child participants must ignore extraneous distractions (testing usually takes place in a noisy and vibrant daycare environment) and follow the experimenter's varied directions. Testing sessions usually run from 15-30 minutes, occasionally longer, and experiments sometimes require multiple sessions over multiple days to finish. Independent of the grammatical issues underlying the design of the study, there are many non-linguistic hurdles a child must manage in order to provide usable and valid data. Having tested hundreds of children using various comprehension methodologies, my own experience suggests that the minimum age for successfully getting children to participate is around three years. This age criterion seems to be relatively accepted and acknowledged in the acquisition field, as most grammatical comprehension studies usually choose children of three or four years as the youngest group tested.[6] The studies included in this dissertation will therefore be limited to children at least three years old.

The second age criterion that governs which age groups to target for study has to do with the grammatical operations under investigation. If a focus of the acquisition study

[6] This is not to say that linguistic comprehension of children younger than three years-old cannot be investigated. If child productions are taken as an indicator of underlying linguistic competence, then elicited and natural production data, often easier to collect from young children than comprehension data, can be used to probe early knowledge concerning raising (see Chapter 4). Recently, new methodological approaches to having children provide grammatical judgments (e.g. Orfitelli and Hyams, 2008) have been used to push even lower the age at which comprehension can be accessed in acquisition studies. For the types of experimental paradigms employed throughout this dissertation (namely sentence-picture matching and truth-value judgment), however, a minimum age of three years does seem appropriate. In many cases, two-year-olds were included during initial pilot testing, but most such children, especially those younger than two-and-a-half years-old, provided data that proved unusable (usually due to failures on attentional or cognitive/linguistic control conditions). [Spoiler alert!] Most of these issues, however, are obviated in the particular case of the grammatical operation under investigation here (i.e. StS raising), since the relevant window for raising acquisition will turn out to be much older than three years of age. Testing two-year-olds would simply highlight that children younger than about six years simply do not possess knowledge of StS raising. The three- and four-year-olds who were tested serve to underline this point nicely; children even younger therefore are simply not needed.

is age of acquisition, one would logically want to cover an age span that includes both children who have yet to acquire the operations of interest, and children who have acquired the operations. This age span varies by the particular grammatical operation of consequence. As discussed in Chapter 1, parameter setting for word order and basic clause structure, for example, is generally in place before the age of three. Verbal passive acquisition, however, is delayed until some time around the age of seven. It does little good then to test only seven-year-olds if interested in the age of acquisition of word order, just as it makes little sense to test only three-year-olds if interested in the age of acquisition of verbal passives. In cases where cognitive development is tied to age development, due to either maturation or age-dependent learning (e.g. infrequent input), the goal is to span an age range large enough to cover the transition from pre-knowledge to successful acquisition.

In the case of StS raising, the relevant age span to consider for experimental inclusion is dictated by theoretical considerations. On those grammatical maturation accounts that link the acquisition of verbal passives and StS raising, one predicts raising to be acquired at the same age as verbal passives. As reviewed in Chapter 1, verbal passives, when adjectival compensatory strategies cannot apply (i.e. in the case of passives with subject experiencer verbs), are generally acquired between the ages of six and seven. If one is to evaluate the validity of such grammatical acquisition theories, six and seven year-old children must be included in any acquisition studies investigating children's comprehension of raising. A researcher, however, would not want to limit himself to this narrow two-year age range. One would do better, especially for a first study, to cast a wider net by including a broader age range. The inclusion of children both younger than six years-old and older than seven years-old has many benefits for a first foray into assessing when children acquire StS raising. One the one hand, allowing a more extensive age range allows for more precise refinement of the developmental curves of raising acquisition. On the other, if raising and passive acquisition are not tied to a common grammatical denominator, then *a priori*, we have no idea whether StS raising should be acquired earlier or later than verbal passives. To be safe, then, it makes sense to include both younger and older children. Further, given an interest in testing for the grammaticality of unraised sentences, for which many of the relevant grammatical theories under consideration make no predictions for age of acquisition, it would be prudent to include children of many ages. Finally, we need a control group old enough to demonstrate that the experimental method captures linguistic judgments (i.e. a control group old enough to do quite well on the experiment; that is, a group at ceiling performance across all test conditions). While many studies use adult test subjects to demonstrate experimental validity, making use of older children introduces many fewer variables into the experimental paradigm.

For the first comprehensive survey of children's raising comprehension, a wide age range will be considered. At the lower end, children as young as three years of age will be included. As mentioned above, children younger than this often produce unreliable data, so three years-old will constitute the age floor for inclusion. As an age ceiling, nine-year-olds will be the oldest children included. If ceiling performance on raising is not obtained in the nine-year-olds, further experimentation can be undertaken to determine the ultimate general age of acquisition. Given that verbal passives are already one of the last syntactic structures to be acquired around seven years of age, it seems reasonable to cap the first experiment at a nine-year-old age limit. Having settled on ages for inclusion, defining age groups and number of child subjects per group is of next concern. To obtain as reasonably valid developmental curves as possible, one would strive for narrow age groups, with many children in each age group. We will limit age groups to a single age year, proceeding in one-year intervals from three years-old to nine years-old, for a total of seven age groups. To facilitate statistical analyses and comparisons between age groups, each group will contain the same number of children. Group size will be set at 10 children per age group (70 children total) as this number is a reasonable compromise between offering valid group and individual subject analyses, and being reasonable in terms of testing resource allocation. Finally, in order to maintain as valid data as possible, we will strive to include only native English-speaking children, from varied socioeconomic backgrounds.

2.1.4 *Previous Work*

It becomes crucial to reflect on what acquisition data already exist that might be of use in planning one's own initial experiments. In this case, a meager amount of research has been directed to children's understanding of StS raising. This is neither a shot at other acquisitionists, nor at the field of language acquisition. Much of what can be studied in first language acquisition has yet to be probed. Given the complexity of human grammars, one really needs a motivating factor to study the acquisition of a particular grammatical operation. In this case, the conflicting predictions of certain maturational acquisition theories provide the impetus for turning to StS raising.

The first investigators to acknowledge these conflicting predictions and embark upon an initial effort to experimentally inquire into those ideas were Karen Froud, Ken Wexler, and Vina Tsakali. Although their research has never been written up, their results and methodological approach served as inspiration for the experiments contained in this chapter. Thus, it behooves the acquisition field to be made aware of this earliest experimental study concerning children's acquisition of raising. These researchers conducted a two-choice sentence-picture matching task testing both raised (4) and unraised (5) sentences.[7,8]

(4) Homer seems to Lisa to be digging a hole.

(5) It seems to Marge that Bart is driving a car.

They tested 33 children in the age range of four to five years. On unraised, expletive sentences with *seem* and an experiencer (e.g. *It seems to Homer that Lisa is eating a sandwich*), the mean correct response rate was around 80%. On raised sentences (e.g. *Lisa seems to Homer to be eating a sandwich*), the mean response rate was around 45%. That is, on raised sentences, children performed at near chance level (where chance performance would be 50%), whereas on unraised sentences the children performed significantly above chance.

These data are both intriguing and rather limiting. They suggest that young children have few difficulties with unraised sentences (contra EARH), and great

[7] Since the basic methodology of Study 1 generally follows theirs, the reader can get an idea of what they did in the section that describes Study 1 (Section 2.2.1).

[8] The inclusion of the experiencer *to*-phrase is discussed in detail later in this chapter. Also, these authors included sentences without the experiencer, but since the focus of this chapter is on comprehension of raising sentences *with* experiencer-phrases, those results will be ignored until Chapter 3.

problems with StS raising at such young ages (as predicted by many of the grammatical accounts under consideration). That said, without detailed analyses, nor any discussion—either of the motivation behind the methodological choices made or of the conclusions drawn from the obtained data—one is left with more questions than answers. Furthermore, the data represent results from relatively few children, whose ages do not overlap the critical age range of interest (six to seven years-old).

Nonetheless, it cannot be stated strongly enough that this seminal study by Froud, Wexler, and Tsakali set the framework for our own Study 1 below. Their ideas of how to construct a sentence-picture matching task to investigate both unraised and raised forms directly inspired our own testing procedures. We attempt to add to their work by being explicit about our methodological and analytical motivations, by including detailed analyses and discussion, by including a larger group of children that covers a larger age range, and by examining in detail the intricate pattern of errors children make.

2.2 Study 1: Raising (over an Experiencer)

2.2.1 *Experimental Design*

Every empirical investigation has to start with a first probe, some first study or experiment. In this case, to begin a detailed exploration of children's comprehension of raising, that first experiment has rather modest goals. In particular, we hope to address the issues raised in Sections 2.1.1-2.1.4. The primary challenge then for the first study is to determine whether children differentially comprehend semantically equivalent, but syntactically distinct unraised and raised sentences. The basic idea is to see if children have more difficulty with a raised sentence like (6) than with its unraised counterpart (7):

(6) The boy seems to be holding a ball.

(7) It seems that the boy is holding a ball.

The difficulty obviously is how to design an experiment that would allow an experimenter to detect the presence of a grammatical deficit involving StS raising (i.e. non-adult StS raising operation) in the child's grammar. In the case of (6), it is easy to imagine that even a child lacking the syntactic operation(s) necessary to compute the raised structure would nonetheless be able to deduce that the proposition being conveyed involves a boy and a 'ball-holding' event (e.g. focusing on the two nouns and semantically salient verb). That is, regardless of the state of the child's grammar, there exists the potential that the semantic "gist" of a sentence can be gleaned from selective attention to particular elements in the input string. As relates to (6) and (7), a valid comprehension experiment requires that the child, for example, be asked to do more than simply select a picture in which a boy is holding a ball (e.g. as opposed to a girl holding a ball, or a boy doing something other than holding a ball). There has to exist some potential for confusion, such that *if* the child's grammar did not allow the adult representation, *then* errors could be elicited. The question therefore is how to introduce the potential for child errors when testing unraised versus raised sentences.

The "solution" hit upon for this first experiment was to introduce a second character (DP) into the test sentences by means of an experiencer *to*-phrase, as in (8) and (9):[9]

(8) It seems to the girl that the boy is holding a ball.

(9) The boy seems to the girl to be holding a ball.

There are now two potential subjects for the embedded predicate. With respect to 8 and 9, the child must rely on his grammar to determine who, the boy or the girl, is holding the ball. This fact is particularly salient in the raising case of 9, where the subject of the embedded infinitival clause is actually the non-adjacent DP in sentence initial position, not the more proximate DP in the *to*-phrase.

Sentences of the forms in (8) and (9) require, at a minimum, that one determine (i) who is performing the action denoted in the embedded clause, and (ii) to whom it appears that the event in the embedded clause is occurring (i.e. who is thinking that the event of the embedded clause is taking place). If a child's grammar lacked syntactic StS raising, it at least seems possible that the child might misanalyze (9), perhaps confusing who is doing the thinking (e.g. interpreting the sentence to mean the boy is holding the ball while thinking about the girl), or misidentifying who is doing the action of the embedded clause (e.g. interpreting the sentence to mean the girl is holding the ball while thinking about the boy), or perhaps even mixing up both aspects (e.g. interpreting the sentence to mean the boy is thinking about the girl holding the ball).

[9] This idea is credited to Froud, Wexler, and Tsakali (unpublished).

This first study attempts to determine whether or not such mixed up interpretations occur for either unraised or raised sentences. In order to do so, two other concerns must also be addressed. First, it is imperative that results are not simply due to children having difficulties with the general task demands, that is, either following the experimenter's instructions or being able to understand the materials (both sentential and artistic) employed in testing. Some sort of (methodological, not syntactic) control condition needs to be used to determine which children, if any, simply could not follow the experimental testing procedures. The sentence structures used for this rudimentary low-level control condition should be fairly simple, and of a type known within the acquisition literature to offer no grammatical challenge to children of the relevant age range under investigation. Second, if the notion of "thinking" is going to play a role in the interpretation of the unraised and raised sentences, some independent control of children's knowledge of thought and mind must be employed that does not involve StS raising verbs, especially given substantial evidence that until around the age of three or four, very young children do indeed have non-adult knowledge of thought (see Wellman, Cross and Watson, 2001 for a general review concerning "Theory of Mind").

Given all of these concerns and stated goals, four sentence structures were investigated in Study 1: transitive active sentences (10), sentences with the verb *think* and a finite embedded clause (11), unraised, expletive-*it* sentences with *seem* (12), and raised sentences with *seem* (13).

(10) Homer is eating a sandwich.

(11) Lisa thinks that Bart is playing an instrument.

(12) It seems to Homer that Marge is pushing a cart.

(13) Homer seems to Maggie to be bowling a ball.

To assess children's comprehension of these sentence types, a two-choice sentence-picture matching task was developed in which children were shown two pictures side-by-side on a laptop computer screen, and were asked to choose the picture best matching the sentence they were read. Answers were logged by the experimenter on the computer before proceeding to the next item. All sentences were read aloud twice before children were allowed to respond. Item presentation was randomized on an individual subject basis. In order to minimize task demands, only four characters (from *The Simpsons* television cartoon), with whom the children were familiarized during the introduction, were used throughout the experiment. Thought bubbles were used to convey the notion of "thinking" for the *think* condition and both *seem* conditions. The notion of thought bubbles was also familiarized in the introduction. Previous research has shown that typically developing children comprehend such pictorial depictions of thinking (Wellman, Hollander, and Schult, 1996).

The active transitive condition involved pictures in which one character interacts with an object. The foil (i.e. incorrect) picture for this condition had a different, non-mentioned character interacting with the same object. For the *think* condition and *seem* conditions, the correct pictures involved one character thinking about another character performing some action. Thus the picture below (Figure 2.2.1) would constitute the correct selection for the following three sentences: *Lisa thinks that Bart is playing an instrument* (think-condition), *It seems to Lisa that Bart is playing an instrument* (unraised-condition), and *Bart seems to Lisa to be playing an instrument* (raised-condition).

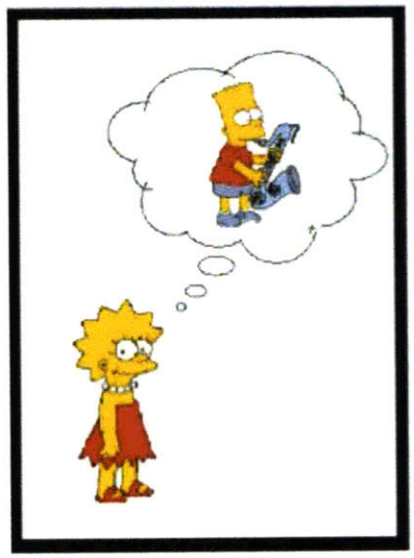


Figure 2.2.1: Example picture in Study 1. Such pictures are paired side-by-side with one of three foil types, and children must choose which of the two pictures best matches the test sentence read to them.

For these three conditions, three different foil types were constructed. *Matrix-reversal* (MR) foils involved switching the character who does the thinking. Thus the MR foil for the picture in Figure 2.2.1 would involve Bart playing the saxophone, thinking about Lisa. *Embedded-reversal* (ER) foils involved switching the character who performs the action denoted by the embedded predicate. With respect to the picture, this would involve Lisa playing the saxophone, thinking about Bart. Finally, *double-reversal* (DR) foils involved switching both who is doing the thinking and who is performing the relevant action. The DR foil to the picture would therefore have Bart thinking about Lisa playing the saxophone. The use of these three foil types allows for the pinpointing of any difficulties in comprehension, whether it is with determining who is doing the thinking (MR foils), with who is performing the action mentioned in the embedded clause (ER foils), or both (DR foils). On any given trial, the child was always presented with the correct picture and one of the three foil types. An example test image is shown in Figure 2.2.2.

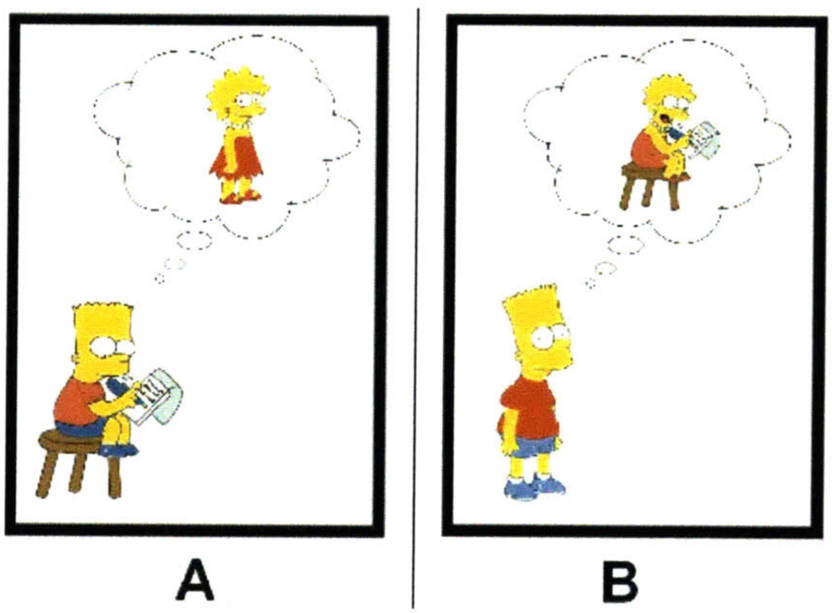


Figure 2.2.2: Example image item in Study 1. For the test sentence *Lisa seems to Bart to be writing a letter* (or *It seems to Bart that Lisa is writing a letter*, or *Bart thinks Lisa is writing a letter*), the picture on the right (B) is the correct response, while an ER foil appears on the left (A).

Each of these foil types was tested six times per condition. Each child thus saw 18 items for the think-condition, unraised-condition, and raised-condition (only 12 items were used for the active-control condition), for a total of 66 test stimuli. Location of the correct picture (left or right side of the screen) was balanced both across conditions and the entire experiment. See the Appendices (A1) for details. Only the raising verb *seem* was used in this experiment, for the reasons mentioned in Section 2.1.1. The use of a single raising verb helped to minimize extraneous testing factors, and served as a parallel to the use of only the single verb *think* as a cognitive control. In any case, the validity of the choice to use the single raising verb *seem* is verified by means of its comprehension in the unraised condition. To the extent that children do comprehend the unraised *seem* sentences, the verb proves to have been a worthwhile choice.

At this point, one must acknowledge the decision to use an experiencer *to*-phrase with the *seem* sentences, both in the unraised and raised conditions. The decision to include the *to*-phrase in this study was based exclusively on experimental, not theoretical considerations.[10] As mentioned previously, it is a real concern that even if children could not grammatically represent raising structures, they might still be able to roughly infer the meaning of a sentence of the form *A seems to be Z* by directly associating *A* and *Z*. How to design a plausible foil for such a sentence using a sentence-picture matching task is not obvious. By including the experiencer, we were able to easily construct an experiment that is simple for children and affords the possibility of detecting difficulties with StS raising. A plethora of concerns regarding the introduction of the experiencer-phrase into the test sentences will be neatly ignored for the present moment. Rest assured, subsequent experimental manipulations will be examined that either manipulate the location of the experiencer-phrase, or do away with it altogether. For now, we examine children's comprehension of StS raising using an experiencer-phrase, knowing full well that any

[10] The many very interesting theoretical considerations introduced by the presence of the experiencer-phrase will be taken up in great detail in Chapter 3.

results obtained from this first study might be limited only to the question of children's comprehension of raising over an experiencer-phrase.[11]

The active transitive sentences are meant to serve as methodological control items to ensure that children were paying attention to the experimental stimuli. The correct items simply depict the mentioned character engaging in the mentioned action. Foils consist of a non-mentioned character engaging in the mentioned action. Given that the correct item could be determined simply by attending to the subject of the sentence, these are a very simple control to assess the most basic level of attentiveness on the part of the participants. There is voluminous evidence that even very young children have no difficulty with such simple subject-predicate active sentences. Any difficulties with this condition would be a reflection of attentional or methodological problems and not core grammar. Thus, those children who did experience difficulties with this condition would be subject to elimination from the experiment, as the rest of their data would be suspect.

The *think* sentences are meant to serve as cognitive controls for the sentences involving *seem*. If children are able to comprehend the *think* sentences, which involve the same pictures and same basic "thinking" notion being tested in the unraised and raised conditions, then any difficulties with *seem* cannot be due to either problems comprehending thought bubbles or a general deficit in theory of mind, since both are shared across the three conditions. On the other hand, data for the *seem* sentences must be judiciously interpreted for those children who do have difficulty with the *think* sentences. That is, if a child is shown to not comprehend the *think* sentences, one would be hard

[11] Which of course would be very interesting in itself. The results, however, hopefully speak to more than just this issue.

pressed to draw strong conclusions about the child's behavior on the unraised and raised conditions, and his knowledge of raising.

Before turning to participant details and the experimental results, considerations should first be given to the predictions of the various grammatical accounts outlined in Section 1.2.6 as applies to the unraised and raised sentences under investigation in Study 1. A theory like EARH, which hypothesizes that premature children find all structures with v_{def} to be ungrammatical (i.e. that all structures must have an external argument), predicts children will perform poorly on all sentences containing a StS raising verb, regardless if the sentence involves StS raising or not. That is, whether or not the embedded subject undergoes raising, the resulting structure (unraised or raised) will involve a projection that lacks an eternal argument, so will be ungrammatical on EARH.

UPR and UFH both predict ungrammaticality for raised sentences, but successful comprehension of unraised sentences. [12] Since StS raising structures demand a defective *v* in adult grammar, on Chomsky's (2001) view, and UPR hypothesizes that premature children's grammar only has fully phasal *v*, raising structures will be ungrammatical for these children. As such, they will either guess at the answer or use a compensatory interpretive strategy, assimilating what would otherwise be an ungrammatical structure to another structure that is grammatical for them. UFH follows an analysis of raising, when an experiencer-phrase is present, where the phrase containing the embedded subject must be Smuggled past the experiencer. If premature children's grammar bans movement from an already moved phrase, that is, does not allow exceptions to Freezing (Wexler and Culicover, 1980; Müller, 1998), then any structure requiring Smuggling (e.g. raising past

[12] These are also the predictions made by ACDH. For reasons discussed in Section 1.2.4, however, ACDH has been shown to be untenable on other empirical grounds (e.g. failure to account for movement of the external argument to subject position).

an experiencer and verbal passives) will be ungrammatical. As such, according to UFH, children are predicted to have comprehension difficulties with the raised sentences tested here, which include an experiencer. Unraised sentences do not require any such syntactic juggling, and therefore will not be problematic for children subject to UFH.

CAH claims that children are unable to represent structures that derive a mismatch between a syntactic position and the canonical theta-role associated with that position. Such an account is compatible with children's early knowledge of unraised sentences, which do not violate any such mapping rules (the embedded subject is in subject position and no theta-role is assigned to an expletive subject). According to CAH, children should also not have any problem with the raised sentences tested here, which also do not violate the aforementioned mapping rules. The raised sentences in Study 1 all involve derived subjects that receive an agent theta-role from the embedded verb.[13]

2.2.2 *Participants*

Motivations for participant inclusion are laid out in Section 2.1.3. Data were ultimately gathered from 70 children (34 girls, 36 boys), with 10 children in every one-year interval from three to nine years of age, with participant details below in Table 2.2.1. All participating children were normally developing native-English learners and came from families of varying socioeconomic status throughout the Boston/Cambridge (MA) area.

[13] It is somewhat unclear how CAH should apply to raised sentences. If children are assumed to know that the subject position of a raising verb is non-thematic, then CAH predicts no mapping rule violation if the embedded subject theta-role is agentive (which in these sentences, it is). Otherwise, as already mentioned, all the bi-clausal test sentences in Study 1 involve agentive embedded clause matrix verbs (see A1 in the Appendices), which means every raised subject will receive an agent theta-role, and the mapping between every subject position and an agent theta-role is maintained. In either case, however, CAH predicts children to have no difficulties with these raised sentences.

Age Group	#	Mean Age	Youngest	Oldest	Male	Female
3	10	3.51	3.04	3.87	3	7
4	10	4.52	4.18	4.95	6	4
5	10	5.49	5.13	5.83	6	4
6	10	6.45	6.03	6.98	7	3
7	10	7.52	7.05	7.79	3	7
8	10	8.45	8.10	8.77	6	4
9	10	9.49	9.05	9.96	5	5
Total	70	6.49	3.04	9.96	36	34

Table 2.2.1: Child participant details for Study 1.

2.2.3 *Results*

The experimental results for all children, collapsing momentarily across foil type, are summarized in Table 2.2.2.

Age Group	Actives	Think	Unraised	Raised
3	100.0%	88.3%	85.6%	43.9%
4	99.2%	92.8%	88.9%	45.6%
5	99.2%	95.6%	92.8%	44.4%
6	99.2%	95.6%	91.7%	51.7%
7	100.0%	96.1%	96.7%	71.1%
8	99.2%	98.3%	98.9%	75.6%
9	100.0%	100.0%	98.9%	92.2%
Average	99.5%	95.2%	93.3%	60.6%

Table 2.2.2: Accuracy for all conditions across all age groups, collapsing across foil types, including data for all children for Study 1.

Overall, children performed extremely well on the active transitive condition. All age groups were 99% accurate for these control trials. No child made more than a single mistake on this condition, and only four of 70 children made even one mistake. Thus, all children performed at a statistically significant above-chance level on this condition, and therefore no children are omitted from subsequent analyses due to inattention. Children likewise generally performed quite well on the *think* trials, with all age groups scoring above 88% correct (detailed analyses of the *think* condition appear below). This indicates that children for the most part have no difficulties either with the concept of thinking or the manner in which it is depicted in this experiment vis-à-vis thought bubbles. Similarly, children performed quite well on the unraised condition, with no age group's mean score below 85% accuracy (detailed analyses of the unraised condition appear below). Children thus generally have no difficulty comprehending the verb *seem* with an experiencer-phrase, at least in its unraised form.

Most children, however, had great difficulty with the raised sentences. Across the 40 youngest children, accuracy did not differ from chance level ($t(39) = -0.978$, $p = 0.334$). No age group scores noticeably better than statistical chance level until the seven-year-olds. Across the first four age groups (three-year-olds to six-year-olds), comprehension as a function of age is flat, with only a 6.7% increase in performance accuracy over this three-year age span. In the subsequent year alone, going from the six year-olds to the seven year-olds, however, performance jumps an impressive 22.8%. This type of rapid growth following years of level stagnation is exactly what is expected on a maturation account, where prior to some genetic event, children lack the necessary grammatical representation to derive the correct sentence meaning, but after maturation such analyses are possible. As discussed in Section 1.2.2, Hirsch and Wexler (2006a, 2006c) note a similar rapid increase in passive comprehension in the seven-year-olds they investigated. This sudden increase in raising performance is further noted in individual subject analyses, counting the number of children in each age group who score statistically significant above-chance (minimum 13 of 18 items correct).[14,15] As seen in Table 2.2.3, before the age of seven, only eight children score at above-chance level on the raised condition. In the subsequent seven year-old group, there are already six children scoring above chance (again, collapsing across foil type, an issue to which we return below). Of the 41 children who fail to score above chance on raising, 78% are less than seven years of age. Meanwhile, 70% of the children seven years-old and older score above chance on this condition.[16]

Age Group	# Children Scoring Above Chance on Raising
3	1
4	2
5	2
6	3
7	6
8	6
9	9

Table 2.2.3: First pass analysis of the number of children scoring at above-chance level for the raising condition in Study 1 (see Footnote 15 for caveats).

[14] This can be calculated by determining a statistical cutoff above which one does not reasonably expect performance if randomly choosing. The equation for above-chance performance is: cutoff = roundup [(items * chance) + (z_{crit} * standard error (SE))]. The SE is calculated as the square root (sqrt) of (items * p[chance for] * q[chance against]). In this particular case, there are 18 items. Chance when randomly guessing between two pictures is 0.5, so p = q = 0.5. Thus, the SE is sqrt(18 * 0.5 * 0.5). Using a 95% one-tailed (only testing above-chance performance) confidence interval, the z_{crit} = 1.65. Calculating the cutoff therefore gives: roundup [9 + (1.65 * 2.12)] = roundup [9 + 3.49] = roundup[12.49] = 13.

[15] Here, above-chance performance on raising is determined by collapsing across foil types and asking whether this aggregate score is statistically above chance. Other statistical approaches, sensitive to foil type effects, are explored further along in this chapter, and are argued to ultimately offer a more nuanced and sensitive measure of comprehension. It should also be noted that the above calculations make no attempt to exclude any children based on their success or failure on the non-raising conditions, something which will be considered in later analyses. The numbers in Table 2.2.3, therefore, should be considered as a gross and simplistic first-pass analysis of children's comprehension of raising.

[16] Data are presented here for all children, without regard to whether or not each child has been shown to comprehend both the *think* sentences and the unraised sentences. Furthermore, above-chance level is calculated by collapsing across all foil types, and not by verifying that above-chance level is reached for each foil type. Table 2.2.3 therefore represents a rather blunt generalization of children's knowledge of raising. Subsequent analyses will address these very issues.

This very preliminary examination of the data lends support to the class of grammatical acquisition theories predicting that structures involving StS raising are delayed, and likewise predicting no delays for unraised sentences. Children comprehend unraised structures involving *seem* (with an experiencer-phrase), but cannot comprehend their semantically equivalent raised counterparts (also with an experiencer-phrase) until around the age of seven. Before this age, very few children (only about 20%) comprehend the sentences involving raising, whereas most of the older children do.. While the above data certainly demonstrate a delay for StS raised sentences with *seem*, certain children did have difficulties with *think* and unraised *seem* sentences, as made clear by an examination of performance as a function of foil type. Such failures must be addressed before attempting a detailed analysis of the performance on raised sentences.

Children do rather well (>75% correct) on all foil types with *think* at all ages. Yet for the *think* sentences, it is also clear that children have the greatest difficulties with the MR foils. Children (collapsing across age groups) are overall 5% worse with MR foils compared to the average of the other two foil types. The fact that younger children score noticeably better on *think* trials with ER and DR foils than on trials with MR foils suggests that ER and DR trials are somehow easier with *think* sentences for younger children. This might be because, when given a sentence of the form *X thinks that Y is doing Z*, even children who do not know the meaning of *think*, but who nonetheless correctly parse the embedded clause, would still be able to correctly reject the ER and DR foils, since both foils (incorrectly) involve a picture in which the subject of the embedded clause is not performing the action denoted by the embedded predicate. When presented with the MR foil, however, children cannot simply look to the embedded clause to determine which picture to choose, since both the correct picture and the MR foil have the subject of the embedded clause performing the action denoted by the embedded predicate, and differ only in who is doing the thinking. In order to consistently choose the correct picture over the MR foil, children must understand that the matrix subject in the *think* sentences denotes the experiencer.

Thus a test of children's knowledge of *think* (the verb, the concept, and the materials used to depict such) is whether or not children score well on those *think* trials involving MR foils. We take "above-chance" performance on *think* sentences with MR foils to be 83% accuracy and greater.[17] As indicated in Table 2.2.4, 14% of all the children fail to meet this level of proficiency. Five of 10 children failing on the MR *think* trials are three-year-olds, and nine of 10 are younger than six years-old. Of the 70 children tested, including the 10 children not performing at above-chance level on the MR *think* items, only two children failed to perform at above-chance level on either the ER or DR *think* trials.[18] Thus difficulties with the *think* trials are for the most part

[17] Strictly speaking, statistically significant above-chance performance for six items, even before compensating for multiple comparisons, requires getting all six items correct: roundup[3 + (1.65 * 1.22)] = 6; see Footnote 14 for details on how to calculate statistically significant above-chance performance. To compensate for children's distractibility, a slightly more liberal cut-off has been used (at least five of six items correct, as opposed to six of six items correct). The motivation for this choice is based on years of working with children in experimental settings. As any brief literature review of language acquisition research will clearly demonstrate, children very rarely demonstrate 100% accuracy for any test condition, regardless of their age or how simple the linguistic material, task, etc. Noise on the part of child test subjects is simply a fact. To demand perfect performance in order to deduce child comprehension is setting oneself up to determine children know nothing about anything. A *very slight* noise correction has thus been introduced in the hopes of obtaining more valuable, and hopefully valid, data analyses. In every analysis in this dissertation where such a slight "correction" has been made, the data were also analyzed with the more conservative principle. In *every* such case, the general data trend held, just often with less impact.

[18] These two children, both four years-old, performed at chance level on DR *think* trials, while performing at above-chance level on MR (and ER) *think* trials. Failure on the DR trials means the two children failed to even correctly parse the embedded clause. How they were able to correctly answer the MR trials then is somewhat of a mystery. One possibility is that they had no idea about the meaning of *think*, and guessed on all *think* trials, just happening to achieve above-chance performance on ER and MR trials while scoring at chance level on DR trials. This is not altogether likely, given the remote statistical probability of this

confined to MR foils. These children either do not comprehend the verb *think*, do not comprehend the pictures used to depict characters thinking (i.e. thought bubbles), or have some failure with Theory of Mind, such that they cannot cognitively represent one character thinking about another.[19] Numerous studies suggest that Theory of Mind develops around the age of four (again, for a review see Wellman, Cross and Watson, 2001), which could account for the difficulties that some children, especially the youngest children, had on the *think* trials. Given great success with thought bubbles in other comprehension experiments with children as young as three (e.g. Hirsch and Wexler, 2006a), it seems probable that the difficulties children had here are due to some problem independent of the pictorial representation used to depict characters thinking about each other. In any case, it would be questionable to draw strong conclusions about sentences involving *seem* for these children who do not score well on the *think* sentences, and they are thus excluded from many of the subsequent analyses.

Age Group	# Children Think-MR < 83%
3	5
4	2[20]

occurring (the probability of this can be determined by multiplication of the independent events' probability, the latter which can be determined from the binomial distribution: .7812 [chance on DR] * .1094 [above chance on ER] * .1094 [above chance on MR] = 0.0093). For the 70 children tested, this means we can expect less than one child (70 * .0093 = 0.65) to demonstrate such a pattern *if every child* was guessing randomly for all *think* trials. Another possibility is that both children *did* know the meaning of *think*, but just happen to have become distracted during two DR trials; that is, this result is just due to unfortunate noise. What is not the case, however, is that the two children know *think* (namely, that the subject is the experiencer), but for whatever reason (e.g. processing, distractibility, etc), ignore the embedded clause. Such a child would be expected to get DR and MR trials correct (since only the correct picture has the correct experiencer), but be at chance level for ER trials (since both pictures contain the correct experiencer). The two children in question, however, were fine on ER trials, while at chance for DR trials. Ultimately, since we cannot determine what led to these two children's response pattern, it seems prudent to just consider them as not having acquired the correct meaning of *think*, and as such, they will be eliminated from many of the subsequent analyses.

[19] Given that each of these children scored at least 11 of 12 correct on the active transitive sentences, it is not the case that they simply are not paying attention at all to the task before them.

[20] See Footnote 18. In all, four of the four-year-olds are taken to not comprehend the *think* sentences.

5	2
6	0
7	1
8	0
9	0
Total	10

Table 2.2.4: Number of children in each age group who are not above chance on the *think* trials with MR foils in Study 1. These children are taken to not comprehend the *think* sentences.

Turning to children's comprehension of the unraised sentences as a function of foil type, for all age groups other than the three-year-olds, unraised sentences are answered above 75% correct for all foil types. Again, however, children have the greatest difficulty with MR foils, scoring 10% worse on the MR foils with unraised *seem* compared to the average of the DR and ER foils. Just as children can answer the *think* sentences with ER or DR foils by merely parsing the embedded clause, so too, can children correctly respond to the unraised sentences with ER and DR foils by doing nothing more than correctly interpreting the embedded clause. In a sentence of the form *It seems to X that Y is doing Z,* children cannot just look to the embedded clause with MR foils, since both the correct picture and MR foil have the subject (Y) of the embedded clause performing the action denoted by the embedded predicate (Z). In order to correctly reject the MR foil with unraised *seem*, children must also comprehend who is doing the thinking (i.e. correctly understand the relationship between *seem* and the *to*-phrase experiencer). In order to understand this relationship, children must comprehend *seem* with an experiencer-phrase. Thus, accurate performance on unraised sentences with MR foils serves as a test of whether or not children comprehend the verb *seem*. Again taking "above-chance" performance to be 83% accuracy or above (minimum 5 of 6 items correct), 19% of children have difficulty with unraised *seem* (see Footnote 17). As shown in Table 2.2.5, all of the children who have trouble with unraised sentences are younger than seven years of age. Fully 60% of the three-year-olds fail on unraised *seem*, constituting roughly half of all those who fail.

Age Group	**# Children Unraised-MR < 83%**
3	6
4	3
5	2
6	2
7	0[21]
8	0
9	0
Total	13

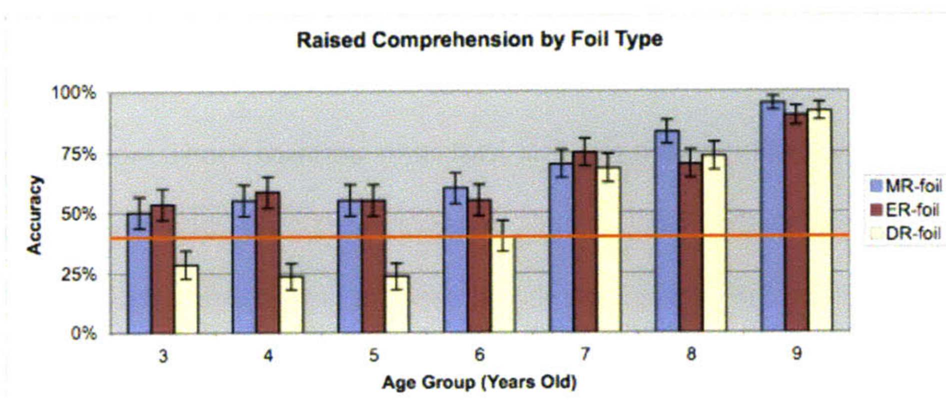
Table 2.2.5: Number of children in each age group who are not above chance on the unraised trials with MR foils in Study 1. These children are taken to not comprehend the unraised sentences.

These data make clear that while the good majority of children four years of age and older comprehend the verb *seem* (at least in its unraised form), most three-year-olds do not. This suggests that if one intends to use unraised forms as an independent lexical control for studying StS raising, one should consider not including three-year-olds, as at

[21] Every child, but one, who failed to comprehend the *think* sentences also failed to comprehend the unraised sentences. This is expected on the assumption that the reason children fail on the *think* trials has to do with either premature Theory of Mind or difficulty with the experiment itself. Such issues should straightforwardly carry over to the unraised trials as well. One seven-year-old, as noted in Table 2.2.4, failed to achieve above-chance performance on MR *think* sentences. This is a bit surprising considering that a typical seven-year-old should find such sentences trivial to comprehend. Further complicating matters is the fact that this child correctly responded to the MR unraised trials, indicating that he comprehends unraised sentences. That this seven-year-old performed well on the unraised sentences, but not on the *think* sentences suggests that he likely did comprehend the *think* trials, but had a momentary lapse in concentration. He correctly answered 16/18 *think* sentences, missing just two MR trials. Still, to be conservative, this child's data will be omitted in many subsequent analyses.

least in this case, the majority do not know the meaning of the raising verb used in its unraised form, which makes asking questions about the raised form dubious at best. The results of those children who do not comprehend *seem* in its unraised form cannot speak with authority to the question of children's comprehension of StS raising, and will be omitted from many subsequent analyses.[22] Three-year-olds, having performed so poorly on this condition as a group, will also be omitted from certain group analyses below.

Children's comprehension of the raising sentences appears quite different from their comprehension of the other sentence types. Considering all 70 children for a moment, all foil types for the raising condition are answered at or below statistical chance level until seven years of age, as seen in Figure 2.2.3 (below-chance / chance level is indicated by the orange line at 39%).[23] As a function of age groups, MR and ER foils are answered near chance level before age seven, while DR foils are consistently answered at below-chance level.



[22] Every child who was found not to comprehend unraised sentences also failed to comprehend raised sentences. This is a nice finding, as it would be difficult to explain how a child could understand the raised sentences but not their unraised variants, at least given the comprehension theories under discussion.

[23] Below-chance performance was calculated for six items per (foil) condition across an age group of 10 children (i.e. 60 items per condition), not correcting for multiple comparisons, which yields a mildly liberal cutoff. [Below chance cutoff = [items * chance] – [zcrit * SE] = (60 * 0.5) – (1.65 * 3.87) = 23.6; as a percentage: 23.6 / 60 = 39%.]

Figure 2.2.3: Comprehension accuracy for all 70 children on raised sentences for the three foil types in Study 1. Error bars represent one standard error, and help substantiate the claim that children younger than six years-old are below chance for DR-foils (since the SE error bars do not overlap the orange chance line at 39%).

Unlike for the *think* and unraised conditions, where children can respond correctly with some foils (i.e. ER and DR foils) simply by interpreting the embedded finite clause, no such possibility exists in the case of the raised sentences. Parsing the infinitival embedded clause (e.g. *to be playing an instrument*) in a raised sentence does not provide any clues as to which picture to choose since it contains no adjacent lexical subject. The non-finite clause itself is not an independent grammatical string. Comprehension requires linking the sentence-initial DP and the embedded clause (i.e. StS raising), while also correctly interpreting the experiencer-phrase. This holds for raised sentences when paired with all three foil types. One cannot simply look to performance with one foil type to deduce whether children comprehend the raised condition, as was the case with the MR foils for the *think* and unraised conditions.

How then to define above-chance performance for raising? In constructing Table 2.2.3, statistically significant above-chance behavior was defined as correctly answering 13 of the 18 raising sentences. While this metric allows one to identify above-chance performance for the condition, it does not require that above-chance performance be observed across each foil type. For example, using this criterion, a child answering all six MR and DR raised sentences correctly, but answering only one ER raised sentencescorrectly (total 13/18 correct) would be considered to have demonstrated above-chance performance on the raised condition. It is not at all obvious, though, that this is the conclusion that one would want to draw in such a case. A more reasonable and altogether more rigorous test is to require that each of the foil types be answered at above-chance level, so that a child would have to get a minimum of five of six ER, MR, and DR raised sentences correct (total minimum 15/18 correct) to be considered as answering the condition at the above-chance level. While this is certainly a more conservative approach to determining statistically significant above-chance performance (needing 15/18 vs. 13/18 items correct), it also appears to be a much more valid and insightful approach. Of course, children must also demonstrate above-chance performance for the *think* and unraised conditions for reasons noted earlier.[24]

Of the ten three-year-olds tested, only four comprehend the unraised sentences. Of these four, only one comprehends the raised sentences (18/18 raised sentences correct). Based on successful answers to the *think* and unraised sentences, this means that one fourth (1/4) of the three-year-olds whose raising data should be considered comprehend raising. Given that so many of the three-year-olds, 60% in all, failed to comprehend the *think* and unraised sentences, however, it could be that this group comprehension rate of 25% for raising is an overinflated estimate for three-year-olds in general. It could be that the vast majority of three-year-olds do not comprehend raising, but by limiting our analysis to those three-year-olds who succeed on *think* and unraised sentences, we are artificially limiting ourselves to atypically precocious three-year-olds who have already undergone some cognitive maturation that enabled them to succeed, both on the control conditions and the raising condition, where typical three-year-olds

[24] As discussed in Footnotes 21 and 22, no child scored above chance on the raised condition having failed either or both the *think* and unraised conditions.

fail. Thus the mature three-year-olds selected in this manner might comprehend raising sentences at a rate significantly higher than of more representative children of this age. As such, in many of the subsequent analyses, the three-year-old group will be omitted.

Turning to the oldest children, of the 20 eight- and nine-year-olds tested, as already noted, all comprehend the *think* and unraised sentences, and as such, all 20 offer data that bears on the question of when raising is acquired. Collapsing across the two oldest age groups reveals that 75% (15/20) of these children comprehend StS raising. The nine-year-olds performed better than the eight-year-olds, with nine of the nine-year-olds scoring above chance on all three foils, while four of the eight-year-olds made at least two mistakes with at least one foil type, mirroring exactly what is reported in Table 2.2.3 where performance on *think* and unraised conditions was not taken into consideration. Those older children who made mistakes with raised sentences tended to be young for their age group (all of the eight-year-olds not above chance for raising were younger than 8.55 years old, while the single nine-year-old to fail raising was 9.14 years old). As a combined group, accuracy on the raised condition, collapsing across the three foil types, was an impressive 83.9%. Clearly then, eight- and nine-year-olds can be taken to have acquired raising. Therefore, children in these two age groups will be omitted from certain of the following analyses since their near-ceiling performance does not bear on the age window in which raising is acquired. By eight years-old, raising is already acquired.

In order to more clearly determine when raising is acquired, we turn to the remaining age groups. Of the 40 four- to seven-year-olds tested, 31 provided data bearing on the acquisition of raising (six four-year-olds, eight five-year-olds, eight six-year-olds, and nine seven-year-olds) because they comprehend *think* and unraised *seem*. Collapsing across foil type, accuracy on the raised condition is extremely flat across the four- to six-year-olds (Table 2.2.6). Furthermore, accuracy for these three age groups hovers right around chance level (50%). At the group level, four-, five-, and six-year-olds have not acquired raising. Turning to the seven-year-olds, however, reveals an impressive jump of nearly 30% in group accuracy for raised sentences. Group performance for the seven-year-olds is at an above-chance level (79.6% correct). These data generally mirror the findings reported in Table 2.2.2, while using better criteria for subject inclusion.[25] Examining the data of individual subjects, only 32.3% (10/31) of the 31 children comprehend the raised sentences (Table 2.2.6). Among the four- to six-year-olds, this number drops to only 18.2% (4/22) who comprehend raising. Of the seven-year-olds, however, 66.7% (6/9) comprehend the raised sentences. The performance jump noted in the accuracy scores for the seven-year-old group is thus also noted in the jump in above-chance scores for this group as well. These data are also comparable to what is noted in Table 2.2.3, but again, make use of better criteria for subject inclusion.[26]

Age Group	**Raising**		**#**
	Acc	**AC**	
4	50.0%	16.7%	6
5	46.5%	12.5%	8
6	51.4%	25.0%	8
7	79.6%	66.7%	9

Table 2.2.6: Accuracy per age group (Acc) and percentage of children per age group who perform above chance (AC) on the raised condition collapsing across foil type,

[25] Table 2.2.2 represents data for all children, irrespective of their performance on *think* or unraised trials.

[26] Table 2.2.3, like Table 2.2.2, represents data for all children, irrespective of their performance on *think* or unraised trials, and uses a less-stringent definition for above-chance comprehension of raising (i.e. not requiring above-chance comprehension across each foil type as is the case in Table 2.2.6).

considering only four- to seven-year-old children with above-chance scores on *think* and unraised trials in Study 1.[27]

While Table 2.2.6 indicates that raising proves difficult for younger children, it offers no insight into what kinds of mistakes children make when trying to interpret raised sentences. An examination of response errors and successes as a function of foil type allows further insights into how children comprehend StS raising. Staying with the 31 four- to seven-year-olds who comprehend the *think* and unraised sentences for the moment, Table 2.2.7 shows that when collapsing across age groups, statistically significant above-chance performance across the three foil types is quite similar. Roughly 35% to 40% of children correctly comprehend the raised sentences when paired with the various foil types. Regardless of which foil is paired with the correct picture, children successfully respond to the raised sentences a minority of the time.

Raised-DR	**Raised-ER**	**Raised-MR**
35.5%	41.9%	38.7%

Table 2.2.7: Percent of children, collapsing across age groups, who perform at above-chance level on each of the three foil types for the raised condition, considering only four- to seven-year-old children with above-chance scores on *think* and unraised trials in Study 1.

As a function of above-chance performance, success rates with the various foils are generally the same across foil type, both when age groups are collapsed (Table 2.2.7)

[27] For accuracy, means were calculated by collapsing across all three foil types. For determining above-chance performance, children were required to perform at above-chance level across all three foil types.

and when looking within individual age groups (Table 2.2.8). What is also clear, though, is that when accuracy is used as the dependent variable, as opposed to percentage of children above chance, performance across the different foils is not consistent within individual age groups (Table 2.2.8). At least for the four- to six-year-olds, accuracy for the DR foils is noticeably less than for either the ER or MR foils. In fact, accuracy for the DR foils appears to be at the statistically significant below-chance level, whereas accuracy for the ER and MR foils appears to hover right around chance level.

Age Group	Raised-DR		Raised-ER		Raised-MR		#
	Acc	**AC**	**Acc**	**AC**	**Acc**	**AC**	
4	30.6%	16.7%	58.3%	33.3%	61.1%	16.7%	6
5	27.1%	12.5%	54.2%	37.5%	58.3%	25.0%	8
6	37.5%	25.0%	56.3%	25.0%	60.4%	37.5%	8
7	79.6%	77.8%	81.5%	66.7%	77.8%	66.7%	9

Table 2.2.8: Accuracy per age group (Acc) and percentage of children per age group who perform above chance (AC) on the raised condition by foil type, considering only four- to seven-year-old children with above-chance scores on *think* and unraised trials in Study 1.

If statistically significant above-chance performance is taken to reflect a minimum of five of six correct items per foil type, then statistically significant below-chance performance would be a maximum of one of six items correct per foil type. Chance performance, thus, involves getting two, three, or four items per foil type correct. To get a better handle on the errors children make on raised sentences as relates to the various foil types, an analysis of how individual subjects respond is necessary. For children who do not score above chance on the raised condition (i.e. above chance for all three foil types with raised sentences), we can ask at what chance level each foil type is answered.

As seen in Table 2.2.9, for raised sentences with either ER or MR foils it is clear that chance performance is noted overwhelmingly more often than below-chance performance. What this means is that when given a raised sentence presented with a choice between the correct picture and either an ER or MR foil, children do not reliably prefer either the correct picture or the foil picture. Rather, children choose between them at chance level, on average 86.4% of the time, while consistently choosing the foil picture on average only 13.6% of the time, collapsing across ER and MR foils in both cases. When presented a raised sentence along with the correct picture and a DR foil, however, children show tremendous preference for the DR picture (in which the experiencer and agent of the embedded clause are both reversed relative to the correct picture). That is, at the group level, children greatly prefer the DR foil over the correct picture, and do not appear to be randomly guessing when a DR foil is paired with a raised sentence, though they do appear to be guessing when given ER and MR foils (ER and MR foils are answered at chance, not below-chance, level).

Chance Type	**Raised-DR**	**Raised-ER**	**Raised-MR**
BC	75.0%	16.7%	10.5%
C	25.0%	83.3%	89.5%

Table 2.2.9: Percent of children, collapsing across age groups, who perform at below-chance and chance level on each of the three foil types for the raised condition, considering only four- to seven-year-old children with above-chance scores on *think* and unraised trials, and not above-chance scores on raised trials in Study 1.

Chance performance with MR and ER foils and systematic preference for DR foils are both further reflected in the response patterns of individual subjects. For this purpose, the children to consider are again those four- to seven-year-olds who comprehend the *think* and unraised conditions. Of these 31 children, only those who did not perform above chance on raising are of relevance to understanding error patterns. If we take a conservative approach to defining "knowing StS raising" and accept above-chance performance to reflect above-chance scores on all three foil types, then only ten of these 31 children comprehend raising, as can be deduced from Table 2.2.6 and Footnote 27. This leaves 21 children under consideration who were not above chance on raising with all three foil types, (five four-year-olds, seven five-year-olds, six six-year-olds, and three seven-year-olds). What patterns of performance are possible as a function of foil types? For each foil type, a child could answer at an above-chance level (minimum 5/6 correct), at a chance level (2-4/6 correct), or at a below-chance level (maximum 1/6 correct). There are thus nine possible relevant response patterns; there are actually ten patterns possible, but a child scoring above chance on all three foil types is considered to comprehend raising, so is not relevant to the question of error patterns.[28] Pure guessing for raised sentences would be expected to produce a binomially distributed pattern centered around chance level (3/6 correct) performance for each foil type.[29]

As noted in Tables 2.2.8 and 2.2.9, group means (whether across age groups or collapsed for all children) seem to suggest a pattern where children score below chance on raised sentences with DR foils, and at chance level with ER and MR foils. One can

---

[28] Total patterns is calculated as the number of combinations with repetitions, or {{(n + k - 1)!} \ {k!(n - 1)!}} = {{3 + 3 - 1} \ {3!(3 - 1)! }} = 10.

[29] Even with random guessing, children are expected to demonstrate above-chance and below-chance performance at a rate given by the binomial distribution with a probability of 0.5, and requiring either 0 or 1 success over 6 draws.

now ask if this pattern is reflected in individual subjects' responses. That is, are individual children responding below chance to DR foils and at chance to ER and MR foils, or is this pattern merely arising by chance at the group levels? Remember, this pattern is just one of nine possible for children who have not acquired raising. Given the relevant binomial distribution, it is also obviously not the expected pattern if one suspects children to be guessing randomly on raised sentences they have been shown not to comprehend. Nonetheless, a staggering 42.9% (9/21) of the four- to seven-year-olds show exactly this pattern of responses. Clearly children are treating the foil types differently, and are not guessing randomly with raised sentences paired with each of the three foil types. This point is most clearly seen in the case of the DR foils. If children were randomly guessing on raised trials with DR foils, one should only expect to find below-chance performance 10.9% of the time (0 or 1 correct responses out of 6 total, given 50% chance for each item). Yet 71.4% (15/21) of children answer the DR foils at below-chance level.[30] With DR foils, only 23.8% (5/21) of children answer at chance level (if performance was truly random, one expects 78.1% given the binomial distribution), with only one case of a child answering the DR foils at above-chance level when ER and MR foils are not also at above-chance level. Meanwhile, children respond to ER foils at chance level 71.4% (15/21) of the time, and MR foils at chance level 81.0% (17/21) of the time (again, 78.1% is expected if children are randomly guessing). Below-chance response rates to ER and MR foils are significantly lower than for DR foils, at 14.3% and

[30] Note that the numbers here are *not* the same as in Table 2.2.9. In Table 2.2.9, for DR foils with raised sentences, 20 children respond at or below-chance level (of 31 total children under consideration). Of these 20 children, 15 score below chance (or 75%). That is, 75% of children who are not above chance on just this (i.e. DR) condition score at below-chance level on this condition. In the textual analysis above, the calculation is different, namely, it asks what percentage of the children not above chance *on all three foils* score below chance on just this (i.e. DR) condition. Here, 15 children score below chance out of the 21 children who did not respond at above-chance level to all three foil types. Similar calculations hold for the other (i.e. non-DR) foil types. Again, these calculations are not those of Table 2.2.9.

9.5%, respectively (10.9% is expected given the binomial distribution). While the response patterns for ER and MR trials look strikingly like what is expected on a model of random guessing, the DR trials are not consistent with guessing. Rather, they reflect child interpretation of raised sentences matching the DR pictures instead of the correct pictures. That these patterns are reflected in individual subject analyses demonstrates that it is individual subject performance that is driving the patterns first noted in the group averages. Possible reasons underlying such patterns of behavior are taken up in Section 2.2.4.

It should be noted, however, that this error pattern of below-chance performance for DR foils- and chance performance for ER and MR foils appears to hold much more strongly in younger children than older ones. Of the 21 children three to six years old who comprehend the *think* and unraised conditions but fail to comprehend the raised condition, an impressive 52.4% (11/21) demonstrate this exact error pattern (again, out of nine possible error patterns). For these younger children, below-chance performance across DR trials with raised sentences is very common, being noted 76.2% (16/21) of the time. Examining the older (seven- to nine-year-old) children, however, reveals a less pronounced preference for choosing DR foils with raised sentences. Of the eight older children who both know *think* and unraised items, but fail on the raised ones, only one child (an eight-year-old) demonstrates this exact pattern of below chance on DR and chance on ER and MR foils (12.5% of older children).

2.2.4 *Discussion*

Study 1 offers much useful data. First, we gain insight into methodological issues related to the study of the acquisition of StS raising. Three-year-old children prove a difficult group to investigate with respect to their knowledge concerning StS raising. These very young children fail to comprehend various control conditions, whether for issues related to Theory of Mind or linguistic complexity, while others have yet to acquire the lexical items needed to test StS raising (i.e. raising verbs used in unraised structures). Children older than seven years, namely the eight- and nine-year-old age groups, show themselves to have already acquired the relevant raising operations, and only serve to establish the low end of the age range at which raising is fully acquired. As such, these older children will not need to be investigated in future studies. It turns out that the age span to concentrate on lies in the ages between four- and seven-years-old, a narrow span that covers the transition for most children from ignorance of StS raising to mastery. For those children having demonstrated acquisition of the *think* and unraised conditions, the raised condition is clearly delayed until seven years of age, as measured either through group accuracy or by means of number of children within the group performing at above-chance level. Raising comprehension sharply increases between years six and seven. While only 19.2% of three- to six-year-olds comprehend raising, 72.4% of seven- to nine-year-olds have acquired it. The timing of acquisition for StS raising matches the age of acquisition for verbal passives, which are likewise acquired by the majority of children around seven years of age. These age effects, at the very least, are a preliminary hint at the possibility that a common grammatical deficit underlies delays in acquisition for both structures.

Concerning children's behavior on the unraised condition, it is clear that the overwhelming majority of children had no difficulties with unraised *seem* sentences that included an experiencer-phrase. As measured by above-chance comprehension of the unraised MR foils, only 18.6% (13/70) of the 70 children tested had difficulties with unraised sentences. Of those children, 76.9% (10/13) could have had trouble with unraised sentences solely due to problems with Theory of Mind (either generally, or just the manner in which thinking was depicted in the experiment), as they also failed to comprehend the *think* items, also measured on the basis of MR foil comprehension. Failure on both unraised and *think* items could also be due to children having difficulties with more complex linguistic items (e.g. those involving clausal subordination). In all, only 4.3% (3/70) of the children tested had difficulties with unraised items that could not be ruled out independently as problems with some more general cognitive issue.

This is not to say that more children would not have shown problems with unraised items had they succeeded on the *think* condition. In fact, it is surprising that performance on unraised items was generally so high (possible grammatical deficits notwithstanding). A lot of lexical learning needs to go into being able to comprehend a verb like *seem*, even in its unraised form. Furthermore, children had to comprehend the included experiencer-phrase. Even with all these issues, the results show that children were quite successful.

These results concerning unraised and raised sentences argue strongly against grammatical accounts that posit acquisition difficulties for unraised sentences. EARH is one such theory. It predicts ungrammaticality for all structures without an external argument (assumed to be due to v_{def}). EARH correctly predicts children's poor performance on StS raising sentences, which lack an eternal argument, but fails to account for the data demonstrating early acquisition of unraised *seem* sentences, which also lack an external argument. Recent data by Kirby (2005), and Kirby and Becker (2007) serve to further confirm the failure of EARH. These researchers examined the natural production transcripts of four children in the CHILDES database as to use of referential (anaphoric and deictic) *it* and expletive *it* (e.g. as used with weather predicates; *It rained.*). While they find referential *it* to be acquired before expletive *it*, expletive *it* is nevertheless acquired quite early. For the four children examined, age of first use of expletive *it* is as follows: 2;6, 1;11, 2;1; and 2;4. These are also roughly the ages around when expletive *there* is acquired. A theory such as EARH that predicts ungrammaticality for verbal projections without an external argument, whether in the form of raising verbs like *seem* (whether unraised or raised) or weather verbs like *rain*, can no longer be maintained on empirical grounds.

Failure on raised sentences, but success on unraised sentences, however, is compatible with other grammatical acquisition theories being investigated here. This pattern of comprehension is exactly as predicted on Wexler's UPR. Raising structures require a defective *v* in adult grammar in order to allow the embedded subject to raise out of the complement of the matrix *v*P, but such movement is not possible if that *v* is phasal (as hypothesized by UPR). Unraised sentences are licit on UPR since even with v^*_{def}, an expletive can be merged as the matrix subject. Hyams and Snyder's UFH also predicts that the raised sentences in Study 1, but not the unraised sentences, will be delayed. Since raising over an experiencer is assumed by them to involve Smuggling the embedded subject past the experiencer, UFH hypothesizes that such sentences will be ungrammatical in premature child grammar. Unraised sentences do not require Smuggling, and therefore will not be problematic for children subject to UFH.

Hyams, Ntelitheos, and Manorohanta's CAH, a hypothesis about children's mapping between the locus of theta-role assignment and eventual landing sites, fails to predict the raising data. CAH claims that children are unable to represent structures that derive a mismatch between a syntactic position and the canonical theta-role associated with that position. While CAH captures children's difficulties with verbal passives (theme in subject position, a mapping rule violation), and is compatible with children's early knowledge of unraised sentences (embedded subject having not undergone movement and expletive matrix subject not receiving a theta-role, neither of which violate any such mapping rules), it falsely expects children to have no problem with raised sentences, which also do not violate the aforementioned mapping rules, since the derived matrix subjects are agentive (due to all the embedded verbs being so). Thus, on empirical grounds, CAH cannot be maintained.

We have also learned that child errors across foil types are not uniform. Children tend to respond randomly (at chance level) for ER and MR trials with raised sentences, but below chance for DR trials with raised sentences. This is again noted both at the group level and for individual children. This pattern was also found to be stronger in the younger (three- to six-year-old) than the older (seven- to nine-year-old) children, again, not just at the age group level, but even with individual children.

According to several grammatical acquisition theories under consideration, children lack the syntactic means necessary to compute StS raising. This inability to mediate the dependency between the matrix subject position and embedded subject position appears quite compatible with children's poor performance on raising trials. What then to make of the much worse and below-chance performance for the DR foils compared to the other foils? The effect of foil type suggests one of two general possibilities. Namely, that children are applying different interpretive strategies when parsing raised sentences, depending on which type of foil is presented along with the correct picture, or that children are applying the same interpretation to all raised sentences without regard to foil type, but that this singular strategy leads to different results for different foil types. It needs to be stressed here that the term "strategy" is not meant to imply any "strategizing" on the part of the children, but merely reflects the non-adult interpretation that children apply to a string for which their grammar cannot or simply does not afford them the adult grammatical possibility. The details, limitations, and reasons for such a strategy used by children in interpreting raised sentences will be explored in depth in subsequent discussion.

The particular data gathered in Study 1 do rule out several interpretive strategies. First, children are not blindly guessing when presented with a raised sentence. Guessing would result in chance performance across all foil types, but children *consistently* choose the incorrect picture when presented with DR foils. Second, children are not simply analyzing the nearest noun to the embedded predicate (i.e. the DP of the experiencer-phrase) as the subject of the embedded clause. This would predict below-chance performance for both ER and DR foils, but children only demonstrate below-chance performance for the DR foils. Third, they are not simply ignoring the matrix subject plus *seem*, and simply parsing the experiencer-phrase as the subject of the embedded clause, as this would again predict below-chance performance for ER and DR foils. Fourth, children are not selecting the DR foils simply because such pictures are somehow inherently more interesting or attractive. As noted earlier, many children actually prefer the MR pictures with *think* and unraised *seem* sentences.[31]

Those children who have not acquired raising, for the most part, appear to be consistently interpreting a StS raised sentence like *Bart seems to Lisa to be kicking a ball* as meaning the equivalent of the *think* condition sentence *Bart thinks Lisa is kicking a ball.* Such an interpretation maps directly to what is depicted in the DR picture, where both the experiencer and agent of the embedded clause have been reversed, and would straightforwardly account for why children choose the DR foil over the correct picture (Figure 2.2.4).

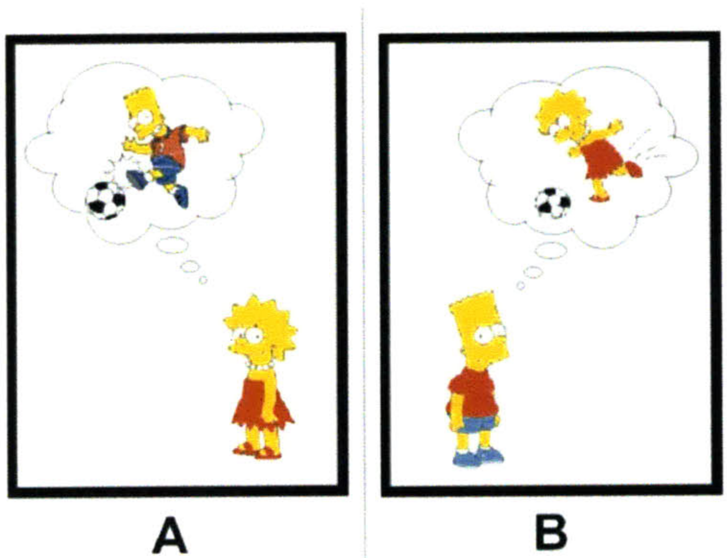


Figure 2.2.4: For StS raised sentences (e.g. *Bart seems to Lisa to be kicking a ball*), younger children consistently choose the DR foil (B) over the correct picture (A),

[31] Note that this also indicates that children's preference for MR foils with *think* and unraised *seem* is not due to any inherent preference for the MR pictures, since with raised *seem* children prefer DR foils.

suggesting they interpret the raised sentence as if it were a *think* sentence (e.g. *Bart thinks Lisa is kicking a ball*).

What about chance performance on the ER and MR foils? Given this analysis, in the case of both the ER and MR foils, neither the correct picture nor the foil match the interpretation. In the case of a raised sentence paired with an ER foil, neither the correct picture nor the ER picture depict the right experiencer if interpreting the raised sentences as *think* sentences. For the MR foils, neither the correct picture nor the MR picture include the correct agent of the embedded clause. In both cases, children are left with an interpretation that fails to match what is shown in either picture, and thus are left to simply guess between the two pictures, which would account for chance performance with ER and MR foil types. That is, (most) premature children appear to consistently apply *one* (incorrect) interpretation to all raised sentences, which happens to result in *different* performance patterns with the various foil types, for the reasons just noted. The particular analysis does not appear to involve children attempting different syntactic parses based on the types of pictures paired with the raised sentences.

While the child errors on the raised sentences are consistent with children responding to raised sentences as if they were the also-tested *think* sentences, it remains unclear if children *grammatically* take a raised sentence to be the same syntactically and semantically as a *think* sentence. In fact, there are reasons to doubt a strict and literal "*seem* as *think*" analysis on the part of premature children, where they are actually taken to substitute the verb *seem* in a raised sentence with *think*. Most notably, such an interpretation would have children ignoring the non-finiteness of the embedded clause of the raised sentence. The English verb *think*, unlike *seem* in the tested raised sentences, is not grammatical with a non-finite embedded clause (14).

(14) *John thinks (to) Mary to dance every Saturday.

While it is logically possible that children take (14) above to be both the meaning of a StS raised sentence with *seem*, and grammatical, this would require a non-adult definition of the subcategorization requirements of the lexical verb *think*, in which case this is *not* a literal *seem*-as-*think* substitution. Another possibility is that children are just ignoring the overt surface syntax and taking the embedded clause to be finite and thus compatible with *think*. While this is certainly possible, it is well known in the acquisition literature that children are sensitive to a finite/non-finite distinction from at least the time of their noted first multiword utterances around 18 months of age (Pierce, 1992a, Poeppel and Wexler, 1993). Why children would be willing to ignore this already established distinction would need to be explained. If an alternative strategy exists that would respect the non-finiteness of the embedded clause, such a strategy might be a better candidate for children's analysis. Perhaps instead children are making use of a linguistic analysis that is both compatible with the *seem*-as-*think* semantics and the surface syntax of the raised sentences.

Before considering what particular syntactic analysis—which we are here calling a "strategy"—that children might be applying in the interpretation of raised sentences, it is worth reflecting on what an interpretative strategy might look like. There are three readily clear requirements for positing a particular strategy. First, the interpretation computed by the strategy must be descriptively accurate. It must capture the empirical evidence from experimental testing. In this case, children's linguistic analysis of raised sentence with an experiencer-phrase must result in the semantics of a *think* sentence. This requires the first DP to be the "thinker" (experiencer) and the second DP to serve as the agent of the action denoted by the embedded clause predicate. Second, the proposed strategy itself must (only) involve syntactic operations allowed by child grammar. There are two subparts to understanding this requirement. For one, the strategy likely cannot involve any great juggling of surface syntax. It seems highly doubtful that children are capable of substituting just any syntactic analysis for strings that they either do not comprehend or for which their grammar cannot provide the adult parse; there is likely a very strong resemblance between the linguistic structure of the strategy and the linguistic form of the string being analyzed. Next, the strategy must not itself be ruled out by the grammatical theory posited to account for the initial ungrammaticality that gives rise to the strategy being invoked. So, for example, if one believed children failed to comprehend raised sentences due to their involving *A*-chains, it makes no sense to posit a child analysis of raised sentences that itself requires *A*-chains. Third, one would like independent empirical evidence that children have actually acquired the particular syntax underlying the strategy. That is, if one thought children interpreted *seem* sentences as *think* sentences, one would want to first demonstrate that children had acquired the syntax of *think* sentences. With these three requirements in mind, we turn now to possible child strategies for dealing with raised sentences.

Given an English sentence of the form DP_1 V DP_2 $TP_{Nonfinite}$ (i.e. the form of the raised sentences in Study 1), the relevant question is how children analyze V, and thus subsequently, the entire sentence to arrive at a meaning similar to that of the *think* sentences; V may be one of the following four types of verbs. First, V could be a StS raising verb, like *seem*, with an experiencer-phrase (15). The second possibility is that the verb could be a subject control verb like *promise,* in which the matrix subject serves as the controller for the subject of the embedded clause (16). Third, it could be an object control verb, such as *persuade*, where the second DP serves as the controller for the subject of the embedded clause (17). Finally, V could be a Raising to Object (RtO) verb, like *imagine* (18). As in the object control case, here too, the second DP serves as the logical subject for the embedded clause. Let us consider each of these four possibilities as a potential candidate for children's analysis of the StS raised sentences in Study 1.

(15) John *seems* to Mary to go to the store every Saturday.

(16) John *promises* Mary to go to the store every Saturday.

(17) John *persuades* Mary to go to the store every Saturday.

(18) John *imagines* Mary to go to the store every Saturday.

Noting the obtained experimental findings, namely the lack of comprehension of StS raised sentences with an experiencer-phrase, we can immediately rule out the possibility that children are (correctly) interpreting the raising verb as such. If children could provide an adult parse of the raised sentences, then they should be error-free in their responses to such sentences in Study 1. The very fact that young children do not comprehend these sentences, but do respond correctly to their unraised counterparts, is enough to dismiss the possibility that children take the verb *seem* to be a raising verb in the StS raised sentences tested here. As such, we can turn to the remaining three possibilities.

Before considering each in turn, however, note that for the three remaining verb types, the presence of *to* before the second DP is somewhat unexpected. Object control and RtO verbs take a second DP that is not part of a prepositional phrase, while for the subject control verb *promise* the second DP is most commonly not part of a prepositional phrase.[32] If children are analyzing *seem* in the raised sentences to belong to one of these three verb classes, then a comment is needed on the presence of *to* in these sentences. One possibility is that children simply ignore it, since the rest of input string would then map onto the syntactic form projected by any of these three types of verbs. This would not appear to be an example of a major juggling of surface syntax; ignoring a one-syllable item in order to obtain a grammatical parse seems decently plausible. Alternatively, children could be taking the *to* in raised sentences as a marker for the very strategy undertaken. Though (most) children appear to have the adult analysis of *seem* in unraised sentences, meaning they correctly interpret the *to*-phrase as picking out the experiencer as indicated by good performance on MR foils, *to* is clearly not serving to mark the experiencer in the raised sentences, since children are not above chance on DR and MR foils. While *to* does not serve to pick out the experiencer in StS raised sentences, it might serve to indicate to the children not to analyze *seem* as the same verb that appears in the unraised sentences. Rather, children might assume that *seem+to* in a raised

[32] While PP "objects" (headed by *to*) are not licit with object control (i) or RtO (ii) verbs, subject control *promise* (iii) is compatible with an optional *to*, though this is certainly less common than the bare DP in natural use:

(i)	John persuaded/convinced/told (*to) Mary to dance.	(Object Control)
(ii)	John imagined/believed/expected (*to) Mary to dance.	(RtO)
(iii)	John promised (to) Mary to dance.	(Subject Control)

sentence should be analyzed as a different verb type entirely; which of the three verb classes that is remains to be determined. Regardless of the exact representational details, it is clear from the acquisition data that some structure with the semantic entailments of a *think* sentence holds for children's analysis of raising structures containing *seem* and a second DP preceded by *to*.

A "*promise*-analysis" for raising, in which *seem* in a raised sentence is analyzed as having the lexical properties associated with the subject control verb *promise*, would share certain traits in common with the adult interpretation of a raised sentence. First, subject control, unlike object control and RtO, associates the matrix DP with the subject of the nonfinite embedded clause, just as is true with raising in the adult grammar, though of course by means of different syntactic operations. Second, as noted in Footnote 32, subject control *promise*, but not object control or RtO verbs, can take *to* before the second DP, as is the case with *seem* with an experiencer-phrase. If children could assume a *promise*-analysis for raised sentences, while they might not know what to make of the thinking event depicted in the experimental pictures, taking the matrix DP as the agent of the action denoted in the embedded clause makes strong predictions about how children might interpret the raised sentences of Study 1. Namely, it suggests that children should select the correct picture when paired with DR and ER foils, since neither of those foils has the matrix subject engaging in the action mentioned in the embedded predicate, as is true for the correct picture. Furthermore, the *promise*-analysis predicts that children should only be at chance for MR foils, as both the correct picture and MR foil have the matrix subject performing the action denoted by the embedded clause. Since children's behavior in Study 1 does *not* conform to these predictions (both MR and ER raised trials are at chance level, and DR raised trials are at below-chance level), it is clear that children do not adopt a subject control *promise*-analysis when hearing StS raised sentences.

One might wonder why this is so if children do lack the syntax of subject-to-subject raising. That is, why do children fail to interpret otherwise ungrammatical raised sentences using the *promise*-analysis (i.e. subject control) given the already mentioned similarities between raising and subject control? There are two clear reasons why children would be expected not to choose a subject control analysis of raising. First, experimental evidence suggests that children do not comprehend *promise* sentences with a second DP like (16) until around the age of seven (Chomsky, 1969).[33] If subject control with an "object" DP is ungrammatical for young children—and the acquisition data strongly suggests that this is so—then a *promise*-analysis for raised sentences would not be available. Second, it is unclear how such an analysis could help children given the test conditions of Study 1. While many verbs can express quasi-similar semantics to *seem* with an experiencer (e.g. *believe*, *imagine*, *understand*), namely associating an experiencer theta-role with one of the arguments, the verb *promise* expresses no such semantics. The act of promising does not involve an experiencer, but rather, an agent.[34] Lacking the notion of "thinking", it is not obvious how children could interpret a *seem*-as-*promise* sentence as mapping to the pictures in Study 1 that involve one character thinking about another. Whether children avoid the *promise*-analysis because they lack

[33] It is interesting to note the similar age of acquisition for both StS raising and such subject control cases. Whether delays for *promise* sentences with an object DP can be related to the syntactic deficits that account for delayed acquisition of verbal passives and StS raising will briefly be explored in Chapter 4.

[34] The theta-role associated with the second DP in a *promise* sentence would not appear to be an experiencer. In (16), *Mary* appears to be the target of the promise, but *Mary* herself need not experience anything at all. That is, there is no psychological state or sensory experience associated with being the target of someone else's promising of something.

the syntax of subject control with an object, or because the verb *promise* does not express the notion of thinking remains unknown. What is clear from the observed data, however, is that indeed children *do* avoid such an analysis.

If children analyzed *seem* in the raised sentences as an object control verb, a sort of "*persuade*-analysis", could this account for their treating raised sentences like *think* sentences? A sentence like (17) involves an agentive matrix subject, and an object that serves as the controller for the agentive (PRO) subject of the embedded clause. As such, a *persuade*-analysis would involve two agent theta-roles, plus the third patient theta-role associated with the object of the object control verb. The pictures employed in Study 1, however, do not involve two agents, since the correct pictures and all foils involve only one character "doing something". The pictures depict scenarios in which one character is performing some action (upon some object, never upon the other character), while a character (possibly the same) "thinks" about the other character. Since the characters themselves never interact, except in the non-physical "thinker-thinkee" relation, it is quite unclear how the patient theta-role of an object control analysis would map to these pictures. Perhaps even more importantly, the *persuade*-analysis, like the *promise*-analysis, does not license an experiencer theta-role.

Since no experiencer theta-role is licensed, children would be left with no means of deducing which character is doing the thinking in the test items. Since children *are* consistently (and incorrectly) treating the matrix subject of the raised sentences as the thinker (as is clear with DR foils), it is unclear how they would come to do so on an object control analysis that does not provide for such an experiencer. Perhaps children using the *persuade*-analysis thus ignore the matrix subject, and just concentrate on the embedded clause. Since object control is in effect, they take the object to be the logical subject for the embedded predicate. Just focusing on this relationship, namely who performs the action of the infinitival clause, predicts children should be at below-chance level for raised sentences with DR and ER foils, and at chance level for MR foils. Since these predictions are not borne out by the experimental findings, children are not analyzing the raised sentences along these lines.

Let us remind ourselves of the experimental findings and of what analysis children are applying in their interpretation of raised sentences. Young children who do not comprehend raised sentences with an experiencer are at chance on ER and MR foils, but below chance on DR foils. Such behavior is consistent with interpreting the raised sentences, regardless of the foil with which hey are paired, as having the same meaning as a *think* sentence: the matrix subject (first DP) is taken to be the experiencer ("thinker"), while the sentence medial experiencer (second DP) is taken to be the agent of the embedded predicate. Such an interpretation naturally maps onto the DR foils, and maps to a meaning that matches neither the correct picture nor the ER and MR foils. Thus, children consistently choose the DR foil over the correct picture, producing below-chance performance, and are left guessing between the correct picture and the ER and MR foils, resulting in chance performance. A literal *think*-as-*seem* analysis, however, was considered undesirable given that it would require children to ignore the non-finiteness of the embedded clause in the test items, since *think* subcategorizes for a finite embedded clause (14).

What if children analyzed *seem* in the raised sentences as a RtO verb, along the lines of (19)?

(19) John imagines/believes/understands Mary to wear a hat (every Sunday).

An *imagine*-analysis would appear to be *very* compatible with how children comprehend the raised sentences. Unlike with the *promise*- and *persuade*-analyses, an analysis based around a RtO verb gets the experiencer theta-role of the matrix subject correct, in the sense that it maps to how children (incorrectly) are found to actually interpret raised sentences in Study 1. For the RtO verbs in (19), not only is the theta-role correct at a general thematic level (e.g. experiencer vs. agent), but it is precisely the semantic notion (one of "thinker") that is being conveyed in the experimental pictorial depictions.[35] Crucially, the "thinker" subject of a verb like *imagine* is basically the same semantic experiencer as the subject of *think* or the experiencer-phrase in a raised sentence. As such, children would have little difficulty mapping the subject of *imagine* to the character doing the thinking in the test pictures. Yet, not only does the *imagine*-analysis get the matrix subject correct, unlike the previously-discussed alternative analyses, but it also gets the logical subject for the embedded clause right. The object of a RtO verb is also the subject for the infinitival embedded clause. Therefore, children subject to such an analysis should take the second DP as the character engaging in the action denoted by the embedded predicate. Taking the first DP as the experiencer, and the second DP as the agent of the embedded action would precisely lead to below-chance performance on raised sentences with DR foils, since the DR foil matches this interpretation. It would also lead to chance performance with ER and MR foils, since

[35] Other RtO verbs also lexically require experiencer subjects, but not necessarily "thinkers". For example, the subject of either *want* or *need* is not semantically the same as the subject of the verbs in (19).

neither of those get both parts correct. Finally, the *imagine*-analysis gets the non-finiteness of the embedded clause exactly right. The complement clause of a RtO verb is always non-finite, as is the embedded clause of the tested raised sentences.[36]

So it would seem then that the *imagine*-analysis is a great candidate for the strategy used by children when interpreting raised *seem* sentences with an experiencer. One must now ask if such an analysis is a possible candidate given children's grammar. The fact that the RtO-based *imagine*-analysis, posited as an alternative parse for what are hypothesized to be ungrammatical StS raising sentences, has the word "raising" in its name might at the very least cause some concern. Three issues will be addressed. First, what exactly is meant by "raising to object", and how does this relate to the syntactic raising in the raised *seem* sentences? Second, can grammatical accounts predicting delayed acquisition for StS raising nonetheless accommodate raising to object? Third, is there any empirical evidence in the acquisition literature demonstrating that children have indeed acquired raising to object by the age at which the *think*-as-*seem* performance is noted for StS raising sentences?

Whether an *imagine* sentence like (20) involves raising has been an active research area in syntax for over 40 years now. A guiding observation is that a sentence like (20), has a paraphrase with a fully inflected finite clause (21).[37] In the latter sentence, where the complement is a finite clause, *Mary* is clearly the subject of the embedded predicate. The question is what syntactic role *Mary* plays in the first sentence with a non-finite complement.

[36] As already discussed, something would need to be added to account for the *to* that precedes the second DP in the raised *seem* sentences, given that such is unexpected for a RtO *imagine* sentence. Two reasonable possibilities are that children simply ignore *to*, or take it as a trigger for the *imagine*-analysis.
[37] Object control sentences do not have a paraphrase with a finite complement clause (i):
(i) *James persuades/tells/asks that Mary is dancing

(20) James imagines Mary to be dancing.

(21) James imagines (that) Mary is dancing.

Certain expressions generally taken to be indicative of subjects can appear in the post-verbal position with an infinitival embedded clause: existential *there* (22), weather *it* (23), and idiom chunks (24).[38]

(22) James imagines there to be ice cream at home.

(23) James imagines it to be raining outside.

(24) James imagines the cat to be out of the bag.

At the same time, the post-verbal DP in such sentences also demonstrates certain properties usually associated with a direct object. If the full DP *Mary* in (20) is replaced with a pronoun, it appears in the (objective) accusative form (25), and not in the nominative form as would be the case if (21) had replaced *Mary* with a pronoun (26).

(25) James imagines her/*she to be dancing.

(26) James imagines (that) she/*her is dancing.

[38] The syntactic observations reported in (20)-(28) are taken from Runner (2006), credited to Rosenbaum (1967) and Bach (1974).

Likewise, if *Mary* in (20) is replaced with an anaphoric element that is coreferential with the matrix subject, it appears as a reflexive (27), not the pronoun that is required with a finite complement clause (28).

(27) James$_i$ imagines himself$_i$/*him$_i$ to be dancing.

(28) James$_i$ imagines (that) he$_i$/*himself$_i$ is dancing.

Given these data, the post-verbal DP in a sentence like (20) has both subject and object properties. Linguists have differed in how to account for this mixed pattern.

Rosenbaum (1967) was the first to propose a raising analysis for such sentences, in which the post-verbal DP starts as an embedded subject that is then taken to raise to object position via transformation. Postal (1974) furthered the arguments in favor of the raising approach, offering over a dozen syntactic tests to support the raising idea. If the post-verbal DP is base generated as the subject of the embedded clause and then raises to object position, the mixed properties noted in (22)-(28) are straightforwardly captured since the DP is at times both a subject and an object.

A different approach is adopted in Chomsky (1981), where the post-verbal DP in (20) is assumed to be in the subject position of the embedded clause at all levels of representation. As such, its subject properties are directly accounted for. As for the object properties, Chomsky claims these are not actually properties of direct objects, but rather, properties that hold of a DP in a particular syntactic relationship with a nearby verb. It just happens that these properties apply to both direct objects and the embedded subject of a sentence like (20), but do not apply to a sentence like (21) where the subject is too "far away" as part of the finite embedded clause. A crucial characteristic of verbs like *imagine*, *expect*, and *believe* when taking a non-finite complement clause is that they "exceptionally" assign case to the post-verbal DP, even though it is not a direct object of the verb. This approach came to be known as the exceptional case marking (ECM) account.

While the ECM approach to sentences like (20) dominated through the 1980s, it has since widely been supplanted by the RtO analysis. This came about due both to new evidence and to reconsideration of old evidence in favor of the RtO approach. Lasnik and Saito (1991) offer several arguments, some novel and others renewed, that the infinitival subject in (20) occupies a higher structural position (at some representational level) than the subject of a tensed complement of the same predicate as in (21).[39] For example, citing Postal (1974), they note that *few students* in (30) has wider scope than in (29). They also observe that the post-verbal DP causes a Binding Theory Condition C effect in (31) but not in (32). In both cases, the infinitival subject must be high enough to c-command a matrix clause dependent to account for the contrast.

(29) The FBI proved that few students were spies.

(30) The FBI proved few students to be spies.

(31) ?*Sue believes him_i to be a genius even more fervently than Bob_i's mother does.

(32) Sue believes he_i is a genius even more fervently than Bob_i's mother does.

[39] This alone certainly does not prove movement to object position, but it does demonstrate a clear structural difference between cases with an infinitival and finite embedded clause.

Also consider another argument from Postal (1974) concerning adverbial interpretation. For an adverb to receive matrix interpretation it is generally assumed to be in the matrix clause. The fact that *incorrectly* can receive a matrix interpretation in (33), but not in (34) where the adverb follows both the matrix verb and post-verbal DP *Greg*, strongly suggests that *Greg* is therefore also in the main clause in (33).

(33) John expected Greg incorrectly to go dancing next Saturday.

(34) #John expected Greg incorrectly would go dancing next Saturday

Further evidence comes from consideration of the "particle construction" (Kayne, 1984). Johnson (1991) points out that if one assumes (uncontroversially) that both "make" and "out" are in the main clause in (35), then *John*, intervening between them, must also be in the main clause.

(35) Mary made John out to be famous.

These data, among other such evidence, have led to a general acceptance of RtO over ECM approaches to analyzing sentences such as (20).[40]

Accepting then the analysis of a sentence like (20) to actually involve raising, we turn next to the question of the compatibility between posited grammatical bans on subject-to-subject raising and the contemporary grammaticality of raising to object (i.e.

[40] Whether raising to object is taken to occur covertly at LF (Lasnik and Saito, 1991, Chomsky, 1995) or overtly in the syntax (Koizumi, 1993, Runner, 1995) is here left unaddressed. In any case, the grammatical acquisition theories under discussion make the same predictions whether this movement is covert or overt.

how can children make use of a grammatical raising structure to interpret a different ungrammatical raising structure?). First, one must not get hung up on the use of the linguistic label "raising". None of the grammatical accounts of delayed acquisition is formulated to work over labels. What matters are the actual grammatical representations and syntactic operations. How linguists choose to refer to such representations and operations is quite irrelevant as far as the analyses are concerned. For example, the formulation of UPR never once references "raising". It just so happens that the syntactic derivation involved in StS raising is banned in child grammar according to UPR. The relevant question then for the grammatical accounts under consideration is whether they predict delayed acquisition for StS raising while still allowing raising to object.

Let us therefore examine each of the relevant theories individually to see whether they might allow the *imagine*-analysis while banning StS raising. According to ACDH, *A*-chains are ungrammatical for premature children. As such, both StS and RtO raising are predicted to be delayed, as both involve *A*-chains in the adult grammar. Therefore, the *imagine*-analysis is not possible under ACDH. Given the already noted problems with ACDH from Chapter 1, however, this is of little concern since ACDH is not under serious consideration as an explanation for the results from Study 1. CAH failed to predict children's difficulties with the raised sentences. Interestingly, CAH predicts that children should find RtO sentences to be ungrammatical, since RtO clearly does result in a mismatch between a syntactic position and the canonical theta-role associated with that position (i.e. agent in object position). Not only should the *imagine*-analysis not be available to children given CAH, neither should any RtO sentences. CAH is thus not a serious contender to explain the StS raising results. On EARH, RtO would be quite acceptable since such sentences do contain an external argument. EARH, however, incorrectly predicts difficulties with unraised sentences, which do lack an external argument, and therefore EARH cannot be entertained as an explanation for the Study 1 results.

Both UPR and UFH correctly predict that children should do well on unraised sentences and perform poorly on (StS) raised sentences. Might they also allow RtO such that the *imagine*-analysis could explain the particular pattern of children's responses to the StS raised sentences in Study 1? Wexler (2004) anticipates the case of RTO and notes that UPR predicts it should not be delayed. To derive (20), the embedded external argument *Mary* first raises to spec,T_{def}.[41] After this cycle, matrix *v* is free to probe and enter an AGREE relation with *Mary*, which is in the same phase so PIC cannot apply to prevent AGREE. As such, the CASE (ACC) feature of *v* is erased, and the derivation (ignoring the matrix subject for brevity) converges. There is no difference between the adult and child derivation of RtO on UPR. Turning to UFH, we also see that RtO should be fine. UFH is concerned with movement from already moved phrases (often requiring Smuggling). Since the adult derivation of RtO does not require such movement, there is no reason children should be delayed (at least for UFH reasons).

While UPR and UFH both ban StS raising (at least with an experiencer-phrase) while allowing RtO, one is still left to ask if evidence exists that indeed children have acquired RtO by the relevant age. Until very recently, no one had examined the acquisition of RtO. Kirby (2009) provides a first look into the acquisition of raising to

[41] Wexler notes that Minimalist Theory argues that the embedded clause of RtO structures involves a defective tense T_{def} and no C. The embedded subject raises to T_{def} and eliminates its uninterpretable features, including EPP, see Chomsky (2001, Footnote 56), but its case-feature is not deleted since T is defective (i.e. does not include a full set of phi-features) and thus remains active.

object, both in production and comprehension. As a first study, she examines all English corpora of the CHILDES database for examples of RtO *want* and *need.* In all, 689 *want* and 19 *need* examples are found to be what she considers very clear cases of RtO.[42] In her second study, Kirby examines 32 children aged four- to five-years-old on their comprehension of simple RtO sentences. For the RtO sentences, four-year-olds scored above 80% correct, while five-year-olds were above 90% correct. These data, while certainly limited, do nevertheless strongly suggest that RtO is acquired early, and certainly well before StS raising. That UPR and UFH both also predict RtO to be acquired early, while ruling out StS raising, along with the specific arguments supporting the *imagine*-analysis, make it a very good candidate for children's interpretation of the raised sentences with a medial experiencer tested in Study 1.

Assuming children are indeed making use of the *imagine*-analysis when interpreting raised sentences, many observations can be made about why children do so. First, and most importantly, given the predictions of the grammatical acquisition theories under discussion, children are unable to provide the (correct) adult derivation for these raising sentences. Nonetheless, children have heard *seem* used many times, most often in contexts where it is clear to them that something like "imagining" is being denoted.[43] At a level of general semantic conception, *imagine* and *seem* share many properties. As the experimental evidence also makes clear, children comprehend sentences with *think* extremely well, probably as soon as they develop Theory of Mind. While children cannot

[42] No age data is given, but going back directly to the corpora shows that many of the RtO *want* sentences come from children as young as two years-old. Kirby (2009) did not look at other RtO verbs, but again, a cursory examination of the corpora demonstrates dozens of examples for various RtO verbs.

[43] As discussed in Chapter 4 (Section 4.1.1), in examining production data from the CHILDES database, it is worthy of note that the utterances children hear containing *seem* might help lead them toward the *imagine*-analysis. The majority of raised sentences that children hear contain animate subjects, where animacy is a prerequisite for sentience, and where only sentient entities may be grammatical subjects of the verb *imagine*. Furthermore, sentences containing *seem* are rather common in the input to children.

comprehend StS raising sentences in an adult manner, they nonetheless attempt to find some reasonable interpretation for such sentences, most likely subconsciously. That is, when parsing the StS raising sentences in Study 1, premature children's grammar is such that the RtO *imagine*-analysis *is* the grammatical parse. It bears noting that the cognitive prerequisites for the *imagine*-analysis are the same as those needed to comprehend the *think* sentences.

It must be asked whether the proposed *imagine*-analysis is a strategy particular to the context of the experiment, or whether it reflects core knowledge on the part of the child. That is, do children merely substitute (a RtO verb like) *imagine*—which is very similar semantically to the also tested verb *think*—for *seem,* given the demands of the experimental task and setting, or do they actually come to the task already having a lexical entry for *seem* along the lines of RtO *imagine*? Also, should we expect the *imagine*-analysis to hold for all (StS) raising verbs?[44] While such an analysis can easily be extended to *appear*, it is unclear how it could apply to certain other raising verbs (e.g. *used* (*to*)). Also, it is important to note that children cannot be extending this analysis to sentences with unraised *seem*, since this would certainly lead to comprehension difficulties, and children performed quite well on unraised structures. It is also interesting to note that younger children who do not know raising employ the *imagine*-analysis more often than older children who do not know raising. It is as if the older children are aware that the *imagine*-analysis does not afford the correct interpretation (that it actually leads to below-chance performance in the case of DR foils), and as such, these children prefer

[44] The examination of child productions containing *seem* in Chapter 4 (Section 4.1.1) shows that while children hear many examples of sentences with *seem*, they produce very few. If children were actively maintaining an interpretation of *seem* as meaning something akin to *imagine*, we might reasonably expect them to produce more utterances with raised *seem*, especially given how frequently some RtO verbs are used by children.

to guess at sentences not licensed by their grammar rather than act on knowingly-wrong meanings. One should also ask how such an analysis would work for raised sentences without an experiencer. These issues regarding the details of the *imagine*-analysis will be taken up in Studies 4 and 5 in Chapter 3. For now, it suffices to note that regardless of the details of the particular strategy employed by premature children, the central take-home point from Study 1 is that raising, at least over an experiencer, is late to be acquired.

The crucial question arising from the discussion of these results, then, is whether or not children's failure to comprehend StS raised sentences is actually due to the grammatical deficits hypothesized by either UPR or UFH, or to a yet unrecognized grammatical problem, or even to no grammatical deficit at all. One wonders, for example, if the difficulties with raising observed in Study 1 could be isolated to the presence of the experiencer-phrase. Certainly the mere presence of an experiencer-phrase alone cannot account for children failing to comprehend a sentence, since the vast majority of children have no trouble interpreting the experiencer-phrase when it appears in an unraised sentence. Could it be that raising over the experiencer, but not raising *per se*, is what causes children's comprehension problems? This is a more difficult question to address, and certainly so if limited only to the data from the current study.

There are a few ways in which raising over the experiencer-phrase could be leading to children's noted difficulties. First, child grammar could simply ban this type of raising. For reasons to be explored in Chapter 3, this is not as stipulative and strange as it might first seem. Given what is known about *A*-movement generally, and raising over experiencers in adult grammar cross-linguistically, there are reasons to wonder if child grammar might not ban such movement. This question can be approached empirically from two directions. First, if children's difficulties with the raised sentences in Study 1 do not follow from the same grammatical deficit implicated in delayed acquisition of verbal passives, then one does not expect that scores on raising over an experiencer and scores on verbal passives will necessarily correlate.[45] The existence of any correlation will be probed later in this chapter in Study 3 (Section 2.4). Second, such a theory would predict that children should have no difficulties with StS raising if the experiencer-phrase was removed. Children's comprehension of raised sentences without an experiencer-phrase will be taken up in two studies in Chapter 3.

Another reason children could have failed to comprehend the raised structures in Study 1 also involves the experiencer-phrase, but not a grammatical ban on raising over it. Rather, it could be that children have a processing, but not grammatical, problem when it comes to linking the displaced subject with the embedded predicate, given the presence of the intervening experiencer-phrase. For reasons to be discussed, theories of adult language processing posit increased processing loads when linking syntactic elements across intervening DPs. If children have more limited processing resources than adults, then perhaps the increased processing loads encountered with raising over an experiencer simply prove too much for their language processors (but not their grammars).[46] A processing account such as this can also be studied via StS raising correlation with verbal passive comprehension (it predicts no necessary correlation) and in StS raising studies

[45] Such a prediction is of course entirely dependent on the particular details of the grammatical acquisition theory under consideration. UFH predicts that verbal passive comprehension and raising comprehension over an experiencer will be correlated in an individual child's grammar, but expects no such correlation for verbal passives and StS raising without an experiencer. UPR, however, expects StS raising and verbal passive comprehension to be correlated regardless of the presence or absence of the experiencer-phrase in StS raised sentences.

[46] This assumes a theory of language processing where the processor and grammar are not one and the same (cf. Phillips, 1996). It also assumes children's processing resources are significantly different from those of adults, which, while possible, is something that data do not currently support (Crain and Wexler, 1999).

without an experiencer-phrase (it predicts no difficulties once the intervening element is removed). A third method for studying such a processing theory is simply to dislocate (e.g. front) the experiencer, so that it no longer serves as an intervener in the final (surface) representation. If errors with raising disappear with a fronted experiencer-phrase, it would seem that processing difficulties, and not grammatical difficulties, are to be implicated in explaining the results from Study 1. It is to this possibility that we turn next.

2.3 Study 2: Raising Comprehension with Fronted Experiencer

2.3.1 *Motivation*

As discussed at the end of Section 2.2.4, there are multiple reasons why the intervening presence of the experiencer *to*-phrase between the derived surface subject and embedded clause might have contributed directly to children's selective difficulties with the StS raised condition in Study 1. First, children's comprehension problems might be due to their grammar ruling out raising over an experiencer. Many languages that have unraised structures with an experiencer-phrase, nevertheless do not allow raised structures with an experiencer-phrase (raised structures without an experiencer-phrase, however, are fine). This possibility of a grammatical ban on raising over an experiencer in child grammar will be taken up in Chapter 3. Second, children might have no grammatical problem with raising, not even for raising over an experiencer-phrase, yet the added processing cost associated with mediating the long-distance dependency between the surface subject and its trace position in the embedded clause might prove too great, and cause children to fail on the subject-to-subject raising sentences.[47]

A processing story along these lines should be considered and addressed both logically and experimentally. By manipulating the surface structure position of the experiencer, we might determine that children in fact do not have any difficulties with StS raising *per se*, just with raising over the experiencer. Of course, manipulating the surface position of the experiencer-phrase might also do little, or even nothing, to affect children's comprehension of raising. This could be for two distinct reasons. First, the position of the experiencer might be irrelevant, where its mere presence is enough to cause a grammaticality problem for children. If this is the case, a different experimental approach is needed to address children's comprehension of raising (e.g. one in which the experiencer is altogether omitted; see Chapter 3). Second, the experiencer might have nothing to do with children's difficulties in Study 1, such that StS raising alone was proving to be the culprit for the noted difficulties with the raised sentences. If so, one expects to find children having trouble comprehending all raising sentences, regardless of the presence or surface position of the experiencer-phrase.

Before turning to experimental means of studying raising with a dislocated experiencer, let us first consider why psycholinguistic research predicts increased processing costs for movement across an intervening DP. Consider one widely referenced model of sentence processing, Gibson's (1998, 2000) Dependency Locality Theory (DLT). DLT identifies two important aspects of memory required for sentence

[47] An explanation would still be needed for the observed *pattern* of results in Study 1, namely better performance on the MR foils with *think* and unraised sentences, and the worse (below chance) performance on the DR foils with raised sentences. This, however, could take the form of children attempting an alternative parse along the lines of the *imagine*-analysis given the hypothetical processing limitations (see Section 2.2.4).

processing: *storage* and *integration*. The storage component is concerned with predicting the (minimal) number of syntactic heads necessary to grammatically complete the current input string, such that increased processing costs are associated with carrying along a greater number of predictions (Gibson, 2000). While storage plays an important role in determining sentence processing costs, it is not the crucial component when trying to understand why the presence of the experiencer might lead to children's comprehension difficulties.[48]

Turning to integration, DLT proposes that the difficulty of integrating a new element into a sentence structure during parsing increases in proportion to the distance between the heads of the two constituents (i.e. integrands) being integrated. One means of measuring this distance relates to decay, where the activation of an attachment site is thought to diminish over time (Warren and Gibson, 2002). Decay is directly related to the number of intervening elements that occur between the integrands.[49] On such an account, in the case of StS raising, the distance between the derived subject and its trace in the embedded clause increases with the inclusion of the experiencer. Processing is therefore going to be easier when integration does not have to take place over the experiencer-phrase, such that the subject being held in working memory will undergo less decay by the time the embedded clause is to be integrated when no intervening experiencer need be

---

[48] On hearing the input fragment (i), DLT predicts a storage cost associated with the need for some predicate (either a bare adjective or a non-finite VP; e.g. *John seems fat*, *John seems to be sleeping*).

(i) *John seems...*

Given a fragment that includes the experiencer, DLT makes a very similar prediction, namely, the need for some predicate.

(ii) *John seems to Mary...*

To the extent that bare adjectives are ungrammatical with raising and an experiencer (e.g. *John seems to Mary *(to be) fat*; see Section 4.1.1), storage requirements should actually *decrease* when the experiencer is included (bare adjectives should no longer be considered) if storage costs are associated with the size of the search space, as opposed to simply the number of syntactic elements. Regardless, it is unclear how storage could be taken to predict *increased* processing when the experiencer is included.

[49] For Gibson (2000), the relevant interveners are taken to be DPs. If decay is merely due to temporal lag, however, any intervening linguistic material should be relevant for increasing integration costs.

processed. The distance metric of DLT can also be quantified in terms of interference, such that the cost of integrating a new element is taken to depend on the degree of similarity between the integrands and any intervening elements (Gordon, Hendrik, and Levine, 2002). This source-memory interference model also predicts increased costs when processing raised sentences with an experiencer, since the subject and experiencer are both proper names in Study 1.[50]

While a processing account like DLT predicts greater processing costs when interpreting sentences with raising over an experiencer versus when there is no intervening experiencer, this by itself hardly suggests children should altogether fail to comprehend raising over an experiencer. First, what is needed is an extra stipulation that children's processing resources are significantly less than those of adults. If such was not the case, then there is no explanation for why adults comprehend raising over an experiencer, given the increased processing costs, versus when the experiencer is not present. That is, to some significant degree, adults would have to have a larger pool of processing resources from which to draw as compared to children in order for the increased difficulty of raising over the experiencer to somehow cause only the children's processor to crash. Second, it is quite unclear how a processing explanation, were it to account for children's difficulties with raising over the experiencer, could account for the particular pattern of responses children give to the raised sentences with the various foil types in Study 1.[51] Third, a processing story is going to have to correctly differentiate

[50] Were the subject and experiencer to be of "different" types under Gordon et al's (2002) definition (e.g. one a proper name, the other a definite description), one would still expect greater processing costs associated with the presence versus absence of the experience, since there would still be some interference between the two DPs compared to no interference whatsoever in the case where the experiencer is omitted.
[51] Of course, it is possible that the particular pattern of results observed with the raised sentences does not itself follow from the processing account, but rather, that given such gross difficulties with processing,

what children can process from what they cannot. One would have to show how raising over an experiencer is too great a burden for the processor, while other complex sentence structures that children are known to comprehend are not equally costly. We will return to this issue in Section 2.3.7.

Given that we do want to examine a processing account, the question then arises of how to manipulate the position of the experiencer-phrase. Both sentence-initial and sentence-final position are grammatical for the experiencer in English, with both raised and unraised sentences (36)-(39).[52]

(36) To the girl, it seems that the boy is holding a ball.

(37) It seems that the boy is holding a ball, to the girl.

(38) To the girl, the boy seems to be holding a ball.

(39) The boy seems to be holding a ball, to the girl.

For the purpose of Study 2, the experiencer will be fronted to sentence-initial position, both for the unraised and raised sentences. This topicalization operation (*A'*-movement) should itself not prove any challenge for children of the relevant ages under investigation. It is well established in the acquisition literature that topicalization, and *A'*-movement in general, is for the most part mastered by the time a child is producing multiword utterances (e.g. see Section 1.2.3 for a brief review). The choice of position

children can do no better than invoke some interpretive strategy. If this were the case, then perhaps children could be adopting the *imagine*-analysis (or some related variant) not due to a grammatical difficulty, but due to a processing one.

[52] One informant found examples with a sentence-final experiencer (e.g. (37) and (39)) to be degraded relative to those with a sentence-initial experiencer (e.g. (36) and (38)). Another seven native English-speakers having been asked, all found both experiencer positions acceptable, while two nonetheless did express a preference for sentence-initial (over sentence-final) positioning. In any case, sentence-initial position of the experiencer, which all eight informants find completely acceptable, is chosen for Study 2.

(sentence-initial or sentence-final) should have no different effect on either the predictions of the various accounts or the results obtained from the children, and sentence-initial position was chosen arbitrarily.

2.3.2 *Changes*

The intention behind Study 2 was basically to replicate the experimental paradigm of Study 1, keeping the overall task as close as possible between studies, while introducing two changes. First, and of obvious importance, was to test unraised and raised sentences with a displaced (i.e. fronted) experiencer. Fronting the experiencer allows one to investigate a multitude of issues related to the raising results obtained in Study 1. If children's comprehension scores with StS raised sentences suddenly improve when the experiencer is fronted, one is more likely to implicate processing as opposed to grammaticality in explaining children's failures in Study 1 on raising over an experiencer. Observing how children respond to the raised sentences with a fronted experiencer will also afford insight into the nature of any interpretive strategies that children might avail themselves of in cases of ungrammaticality. In particular, it will be worth noting how a strategy such as the *imagine*-analysis (Section 2.2.4) could apply in raised sentences with a fronted-experiencer.

Second, Study 2 makes use of different inclusionary criteria for participants, resulting in a slightly different group of children compared to that investigated in Study 1. In particular, comprehension will only be explored with children aged four to seven years. The elimination of the three-year-olds is due to their generally poor performance in Study 1 on the *think* and unraised conditions. In all, 60% of the three-year-olds had to be eliminated in that study. Those results suggest that slightly older children are more likely to offer more useful data with regard to raising, since at very young ages children are still acquiring Theory of Mind, as well as learning which verbs are raising verbs. Not bothering to test three-year-olds is really an attempt to make better use of limited testing resources. Another choice aimed at maximizing testing efficiency is to not bother testing eight- and nine-year-olds. These older children have all basically acquired raising (75% across both ages; 90% for the nine-year-olds) and as such offer little value in helping determine how and when raising is acquired.

Furthermore, only children who are found to comprehend the active transitive, *think*, and unraised sentences will be included for the subsequent analyses for raising comprehension. As discussed for Study 1, without using the *think* and unraised conditions as cognitive and lexical baselines, interpreting results on the raised condition is greatly complicated. By only including children who succeed on these controls, results on raised sentences can be more straightforwardly understood. Those children who fail any of the three non-raised conditions will be replaced as needed to achieve the desired number of children per age group (replacement procedure discussed below in Section 2.3.5). Finally, in order to increase within-group reliability, the number of children per age group will be increased from the 10 children in Study 1 to 15 children for Study 2.

2.3.3 *Strategies*

In case children *do* have difficulties with the StS raised sentences with a fronted experiencer, it remains unclear how comprehension results will vary with respect to the different foils. Will children be at chance for all of them? Will they show above or below-chance performance for some or all foil types? How can or will the *imagine*-analysis apply when the experiencer is not in canonical medial position?

Assuming for the moment that children do lack StS raising syntax, we turn to the question of how the *imagine*-analysis might apply to a raised sentence with a fronted experiencer (40). If children's lexical definition of *seem*, in a raised sentence, was literally something like *imagine*, then the *imagine*-analysis produces a slightly less natural parse (41).

(40) To Mary, John seems to go to the store every Saturday.

(41) ?Mary, John imagines to go to the store every Saturday.

The post-verbal DP in a RtO sentence can be topicalized (i.e. undergo *A'*-movement), but the resulting sentence (41) sounds slightly less natural than a raised sentence with fronted experiencer (40). The RtO sentence in (41) is also noticeably less natural than when the subject is not fronted (20), while no such difference is noted for fronting the experiencer in (40) compared to a raised sentence with the experiencer in sentence medial position (15). It is therefore unclear how fronting the second DP would affect a possible *imagine*-analysis. If children lack the syntax underlying raising, but find a topicalized RtO sentence like (41) to be fine, we predict children should perform exactly the same as in Study 1. Namely, they should interpret (40) along the lines of (41), and as such consistently choose DR foils, but show chance performance for ER and MR foils, for the same reasons already outlined in Section 2.2.4.

If children, however, are unable to entertain (41) for (40), perhaps related to the unnaturalness of (41) for adults, then no s-homophone exists as a possible substitute for a raised sentence with fronted experiencer (40).[53] There are no other possible syntactic structures of the form DP_1 DP_2 V $TP_{Nonfinite}$ that would be grammatical for the children.[54] It would therefore appear that the *imagine*-analysis of raising would offer children no grammatical interpretation once the experiencer-phrase is pre-posed. As such, if premature children nonetheless relied on the *imagine*-analysis, they should respond at chance level across all three foils.[55] Without a convergent parse, they can be expected to do no better than guess randomly. If children, however, are not tied to the *imagine*-analysis, in any form, such that they have not defined StS raising verbs as RtO verbs, then perhaps new interpretive strategies will be discovered when children attempt to interpret StS raised sentences with a fronted experiencer.

2.3.4 *Experimental Design*

Just as in Study 1, children's comprehension of four sentence structures was investigated: active transitive sentences (42), sentences with the verb *think* and a finite

[53] Whether RtO with topicalized infinitival subject is grammatical/acceptable to young children is of course simply an empirical issue in need of determination. If we accept Kirby's (2009) conclusions that RtO is early in acquisition, there is no readily apparent *grammatical* reason why children should not comprehend (41). Whether processing, canonicity, or some other orthogonal issues rules it out, however, would need to be empirically determined.

[54] Once again, both subject control (i) and object control (ii) analyses are grammatical for adults, but almost certainly not possible strategies for children given the arguments in Section 2.2.4, which also apply in these cases where the second DP is topicalized.

(i) Mary, John promises to go to the store every Saturday.
(ii) Mary, John persuades to go to the store every Saturday.

[55] Since it is the topicalized infinitival subject that is awkward in (41), there is also the possibility that children might ignore it, and just try to parse the remaining material (i).

(i) John imagines to go to the store every Sunday.

While this is an ungrammatical string (at least with the verb *imagine*), if children nonetheless make use of it to associate the remaining DP with the infinitival clause, then one can draw conclusions about how children might respond. Just such a possibility will be explored below in Section 2.3.7.

embedded clause (43), unraised, expletive-*it* sentences with *seem* and a fronted experiencer-phrase (44), and raised sentences with *seem* and a fronted experiencer-phrase (45).

(42) Lisa is eating a sandwich.

(43) Lisa thinks that Bart is kicking a ball.

(44) To Lisa, it seems that Bart is digging a hole.

(45) To Lisa, Bart seems to be lifting a rock.

This study makes use of the same sentence-picture matching task as Study 1, using the same pictures and the same methodological procedures (for details, see Section 2.2.1). Children are basically asked to choose which of two pictures best matches the test sentence they hear. Four conditions are again tested, differing only from those of Study 1 in having fronted experiencers for the unraised and raised conditions. First, 12 active transitive sentences are tested as attentional controls. Second, 18 *think* sentences with finite embedded clauses are queried, one-third paired with each of the three foil types, as a sort of cognitive control to ensure that children have acquired Theory of Mind. Third, 18 unraised sentences with the raising verb *seem* and a fronted experiencer-phrase are included, six of them paired with each of the three types of foils. The unraised sentences are meant to serve as a lexical control for the raised sentences, since comprehension of these requires that the raising verb has been learned (at least in its unraised form). Furthermore, they offer a means of ensuring that children do not have a more general problem with fronted experiencer-phrases. Successful comprehension of these unraised sentences would demonstrate that fronted experiencers by themselves cause children no difficulties. Finally, there are 18 raised sentences with the verb *seem* and a fronted experiencer, one-third being paired with each foil type. In all, children will have to respond to 66 items across the experiment. See Appendices (A2) for details. As already discussed, participants will be required to demonstrate above-chance comprehension for all three non-raised conditions in order for them to be included in the final analyses.

Turning to the predictions for the raised sentences with fronted experiencer, the processing and grammatical accounts differ as to whether they expect children to fail or succeed. The processing theory claims that children's difficulties with raising over the medial experiencer in Study 1 was due to the increased processing load (e.g. integration cost on DLT) associated with establishing a long-distance dependency over the experiencer between the raised subject and its trace position within the infinitival embedded clause. If the experiencer-phrase were either removed, or moved so that it did not serve as an intervener in the surface representation, there should be no processing overload and children should be expected to comprehend the raised sentences.

Both UPR and UFH claim that children's inability to comprehend the raised sentences of Study 1 follows from bans in their grammar on particular syntactic operations necessary to compute StS raising. As such, the location of the experiencer-phrase, whether in sentence-initial or sentence-medial position, is irrelevant to whether children will succeed or fail in their comprehension of raised sentences.[56] In the case of

[56] Collins (2005b), upon whose analysis of raising UFH is based, proposes that Smuggling is not required when no experiencer is present, but does not address whether Smuggling is required when the experiencer is topicalized. UFH only predicts comprehension problems in cases where Smuggling takes place. The most logical reading of Collins, however, would suggest that in cases of raising with a fronted experiencer, Smuggling takes place prior to topicalization of the experiencer. As such, UFH, just like UPR, predicts delayed acquisition for raising with a pre-posed experiencer-phrase.

UPR, StS raising alone is enough to ensure failure, while on UFH, the mere presence of the experiencer is enough to see children fail. If children do lack the syntax necessary to compute raising structures, then they should fail on the raised condition in Study 2. If children attempt an *imagine*-analysis with the raised sentences with a fronted experiencer, then children should be at chance for all three foils, since this particular analysis does not produce an interpretation consistent with any of the pictures in Study 2, as described in Section 2.3.3.

2.3.5 *Participants*

The goal for this study was to obtain usable data from 15 children in each one-year age interval from four to seven. To be included, children had to demonstrate statistically significant above-chance performance on the following three tested conditions: active transitive sentences, *think* sentences, and unraised sentences. Above-chance performance for the *think* and unraised sentences was determined by above-chance comprehension of the MR foils only, since success with other foil types is possible even if a child did not understand the sentence (as explained in Section 2.3.2). In order to ensure ending up with the necessary number of children, therefore, more than 60 had to be tested, as certain children inevitably failed to meet above-chance levels on all but the raised conditions. Every child who succeeded on all three of the non-raised conditions had his data included for subsequent analyses. Any child who failed on any of the three non-raised conditions was immediately excluded from the study, and another child of the same age year was tested as a replacement. If that child also failed one of the three relevant inclusionary conditions, he, too, was excluded, and a replacement sought.

This process continued until the requisite number of children was reached for each age group. While the number and ages of excluded children were not formally tracked, children were excluded in roughly the proportions matching the failure rates on the *think* and unraised conditions as observed in Study 1. With replacements, data from 60 children (38 girls, 22 boys) were obtained, with 15 children as planned in every age group, and participant details appearing in Table 2.3.1. All included children were normally developing native-English learners and came from families of varying socioeconomic status.

Age Group	#	Mean Age	Youngest	Oldest	Male	Female
4	15	4.59	4.25	4.90	10	5
5	15	5.49	5.01	5.95	4	11
6	15	6.46	6.08	6.88	2	13
7	15	7.53	7.04	7.98	6	9
Total	60	6.02	4.25	7.98	22	38

Table 2.3.1: Child participant details for Study 2.

2.3.6 *Results*

The experimental results for all children, collapsing momentarily across foil type, are summarized in Table 2.3.2.

Age Group	Transitive	Think	Unraised	Raised
4	98.3%	96.3%	95.2%	74.4%
5	98.3%	97.8%	97.4%	78.1%
6	98.9%	97.8%	96.7%	80.7%
7	98.9%	98.1%	97.4%	91.8%

Table 2.3.2: Accuracy for all conditions across all age groups, collapsing across foil types for Study 2.

As expected, children across all age groups perform brilliantly on the active transitive sentences, the *think* sentences, and the unraised *seem* sentences with fronted experiencers. On the one hand, this is hardly surprising given the strict inclusionary criteria, such that any child not performing at above-chance level for one of these conditions would not have his data included in this analysis. That said, as mentioned in Section 2.3.5, relatively few (less than a dozen) children had to be replaced due to such failures. In any case, it was not hard to find children who performed very well on these first three conditions.

What is perhaps surprising, and at the very least noteworthy, is children's generally decent performance on the raised *seem* sentences with a fronted experiencer. Comparing accuracy for the raised sentences in Studies 1 and 2, collapsing across foil type, reveals that children perform noticeably better in Study 2 (Table 2.3.3). Even among the four-year-olds in Study 2, group performance is right around the 75% mark, compared with 50% comprehension for children of the same age in Study 1. While the largest jump in performance is again noted between the six- and seven-year-old groups, it is much smaller in Study 2 (11.1%) than in Study 1 (28.2%). Do these data, therefore, support the idea that children have no difficulties with raising with a fronted experiencer, and that the difficulties with raising found in Study 1 were merely due to raising over the experiencer-phrase?

Age Group	Study 1	Study 2
4	50.0%	74.4%
5	46.5%	78.1%

6	51.4%	80.7%
7	79.6%	91.8%

Table 2.3.3: Accuracy for the raised sentence condition across all age groups, collapsing across foil types for Study 1 and Study 2.

Such a conclusion seems rather premature given the lack of uniformity and successful comprehension of raised sentences with a fronted experiencer across foil types, as seen in accuracy scores for the various age groups (Table 2.3.4). What these data demonstrate is that while young children are generally quite capable of comprehending the raised sentences with DR and ER foils, much better than was the case in Study 1, they do have great difficulties with the MR foils, for all groups except the seven-year-olds. While on average the four- to six-year-olds accurately respond to the raised sentences with DR foils and ER foils 85.9% and 87.8% of the time, respectively, they only manage to answer the raised sentences with the MR foils 59.6% of the time. Comprehension for these younger children therefore does not differ noticeably from chance performance. From four to six years of age, comprehension for the MR foils is relatively flat (not even an 8% increase in accuracy over the three years). There is, however, a very appreciable jump in comprehension in the seven-year-olds (nearly 25%).

Age Group	**R-DR**	**R-ER**	**R-MR**
4	83.3%	83.3%	56.7%
5	85.6%	91.1%	57.8%
6	88.9%	88.9%	64.4%
7	95.6%	92.2%	87.8%

Table 2.3.4: Accuracy for each age group on the different foil types for the raised sentences in Study 2.

Children's lack of success on raised sentences with MR foils is also apparent in an analysis examining how many children in each age group score at an above-chance level on the raised sentences, broken down by foil type (Table 2.3.5).[57] Even with the four-year-olds, most children successfully answered the raised sentences with DR and ER foils. Again, however, performance on the MR foils is extremely poor. The seven-year-old group is the first in which a majority of children correctly answer raised sentences with MR foils. On average, only 31.1% of the four- to six-year-olds comprehend raised sentences with a fronted experiencer given performance on the MR foils. Again, a noticeable increase in success is noted for the seven-year-olds, 80% of whom manage to answer these sentences correctly with MR foils.

Age Group	R-DR	R-ER	R-MR
4	73.3%	80.0%	20.0%
5	80.0%	86.7%	26.7%
6	86.7%	86.7%	46.7%
7	100.0%	93.3%	80.0%

Table 2.3.5: Percentage of children in each age group performing at above-chance level on the different foil types for the raised sentence condition in Study 2.

It is also worth stressing that while children's performance on raised sentences with DR and ER foils is appreciably better than their performance on the same sentences with MR foils, children are hardly demonstrating perfect raised comprehension with DR and ER foils. Remember, every single child (100%) included in these analyses is

[57] Once again, as was the case in Study 1, above-chance level for the purpose of foil type analyses is taken to be at least 5 of 6 items correct.

demonstrating above-chance performance for *think* and unraised sentences with these very same foils. Yet 90% above-chance levels are not reached for the DR and ER foils with raised sentences until the group of seven-year-olds. And while accuracy scores for *think* and unraised sentences with DR and ER foils hover right around 100%, raised sentences are not answered at above 90% accuracy for both foil types except by the seven-year-olds.

It therefore appears that children younger than seven continue to have greater difficulty comprehending raised sentences than their semantically equivalent unraised counterparts, even with a fronted experiencer. It also seems that for some reason, to be explored below, performance with MR foils is the correct barometer for determining children's difficulties with StS raised sentences in Study 2. Of the 60 children tested, nine demonstrate lack of above-chance performance for raised sentences with DR foils, while eight children are not above chance with ER foils with raised sentences. Fully 34 of the 60 children tested (56.7%), however, fail to comprehend the raised sentences with MR foils.

An individual subject analysis examining whether a child scored at statistically significant below-chance, chance, or above-chance level for raised sentences with a fronted experiencer with MR foils appears in Table 2.3.6.[58]

Age Group	**Chance Type**	**#**
4	BC	2
	C	10
	AC	3
	BC	2

[58] Below chance is at most one item correct, chance is between two and four items correct, and above chance is at least one item correct, out of the six raised sentences with MR foils.

5	C	9
	AC	4
6	BC	1
	C	7
	AC	7
7	BC	0
	C	3
	AC	12

Table 2.3.6: Breakdown of chance performance for MR foil on the raised sentence condition across age groups in Study 2 (BC=below chance, C=chance, AC=above chance). The vast majority of children who are not AC respond at C level, not BC level.

This table again reflects that above-chance level for the majority of children is not reached until the seven-year-old group. Importantly, however, it also demonstrates that in the vast majority of cases, failing to achieve above-chance performance is due to chance, not below-chance, behavior. For those children not scoring above chance on raised sentences paired with a MR foil, 85.3% (29/34) respond at chance level. Only 14.7% (5/34) demonstrate below-chance response rates. It would appear that children who do not comprehend raising are guessing randomly when raised sentences are paired with MR foils. This is supported by the fact that 10.9% of children are expected to score below chance given random guessing on these sentences according to the binomial distribution, which is very close to the actual percentage of children observed to be below chance.

2.3.7 *Discussion*

Overall, child performance on raised sentences, as measured by overall accuracy collapsing across all foil types, is higher in Study 2 than Study 1 (Table 2.3.3). That is, ignoring foil type effects for the moment, children appear to better comprehend raised sentences with a fronted experiencer (Study 2) than raised sentences with a canonical medial experiencer (Study 1). This seems in line with the predictions of a processing account that attributes children's difficulties with raising over an experiencer in Study 1 to the extra processing required to mediate the long-distance dependency over the intervening experiencer. Once the intervening DP is displaced, in this case fronted, comprehension scores should increase. A careful investigation of how accuracy varies by foil type, however, reveals that children continue to have trouble with raising when the relevant subcondition (MR) is considered. The majority of children tested, and nearly 70% of the children younger than seven, fail to demonstrate above-chance comprehension for the raised condition as determined by performance on raised sentences with MR foils.

In fact, performance as measured by the percentage of children above chance (for all three foils) is remarkably similar in the raised conditions across the two studies (Table 2.3.7).[59] Notably, in both studies, a majority of children within an age group comprehending raising is only achieved with the seven-year-olds. The pattern and age of acquisition is therefore very similar across both studies, and again, grossly matches that found in previous studies for verbal passives.

Age Group	Study 1		Study 2	
	Acc	AC	Acc	AC
4	50.0%	16.7%	74.4%	20.0%
5	46.5%	12.5%	78.1%	26.7%
6	51.4%	25.0%	80.7%	40.0%

[59] Data from Study 1 only include those children scoring above chance on *think* and unraised conditions, since such performance was prerequisite for inclusion in Study 2. This allows for better direct comparison between studies.

7	79.6%	66.7%	91.8%	80.0%

Table 2.3.7: Accuracy per age group (Acc) and percentage of children per age group who perform above chance (AC) on the raised condition in Study 1 vs. Study 2.

On the basis of the data from Study 2, one can rule out an account of the results obtained in Study 1 that would attempt to attribute poor performance on the raised condition to children being unable to process the dependency between the matrix and embedded subject position across the intervening experiencer. As measured by above-chance performance on MR foils, children fail to comprehend raised sentences with a non-intervening fronted experiencer roughly to the same degree as they failed to comprehend raised sentences with a medial experiencer. While processing theories suggest that establishing dependencies across intervening linguistic elements results in increased processing costs, in the case of raising for children these costs do not seem to account for delayed acquisition. The surface position of the experiencer-phrase does not determine whether children fail or succeed in their attempts to correctly interpret raised sentences with an experiencer. Whether the mere presence of the experiencer, however, plays some role in the noted difficulties in Studies 1 and 2, remains unclear, at least without further experimentation.

A few more comments are deserved on the issue of processing theories attempting to account for these results bearing on child comprehension of raising. While processing accounts like DLT certainly do predict greater integration costs for the raised sentences in Study 1 compared to Study 2, it is strikingly unclear how any such processing costs could account for children's inability to comprehend the raised sentences in Study 1, given that they successfully comprehend many other sentences requiring syntactic dependencies that cross an intervening DP. It is well known that children have no difficulties producing or comprehending simple object-extracted *wh*-question like (46) well before the ages at which children appear to acquire raising (Stromswold, 1995, Hirsch and Hartman, 2006a; for a review see Guasti, 2002).

(46) Who(m) did Mary kiss?

The raised sentences, however, require extracting an element (the subject) from an embedded clause into a higher clause. While children comprehend sentences like (46) that only require movement within the domain of a single clause, perhaps children's difficulties with raising are related to extracting an element across a clause boundary, where crossing another intervening element might or might not contribute to comprehension problems. As young children have no problem comprehending *wh*-questions where the extracted element comes from the embedded clause like in (47), that, too, cannot be the correct account for noted delays with raised sentences (de Villiers, Roeper, and Vainikka, 1990, Thornton and Crain, 1994).[60]

(47) Who did Mary think was dancing?

---

[60] Production data, both from elicitation and natural setting studies, for long-distance *wh*-questions are slightly more complex, due to the fact that some children produce "medial-*wh* questions" like (i), as noted for English (Thornton, 1990) and Dutch (Kempen, 1994).

(i) What do you think <u>what</u> Cookie Monster eats? [Katie, aged 5;5] (Thornton, 1990)

Such data have not been taken to indicate that children lack the syntax of long-distance *wh*-questions, especially given the comprehension data demonstrating early mastery. Thornton (1990), for example, attributes such cases to some children assuming that agreement in intermediate CPs is required (as in Irish, see McCloskey, 1990), where the intermediate *wh*-word is the overt realization for the spec-head agreement in CP.

Further evidence for children's successful comprehension of sentences involving dependencies across intervening DPs comes from research into the acquisition of relative clauses. Some early studies, such as Sheldon (1974), claimed to find (mixed) delays for object-extracted relative clauses compared to subject-extracted relative clauses.[61] Later studies demonstrate that once felicity conditions on the use of restrictive relative clauses (Hamburger and Crain, 1982) and relevant orthogonal processing concerns (Goodluck and Tavakolian, 1982) are taken into account, all relative clauses, including object relatives, are acquired no later than three years of age. Good performance for object-extracted relative clauses has also been found in a recent study by Hirsch and Wexler (In preparation). Reviewing the literature on relative clause acquisition, Guasti (2002) writes: "[L]ater studies revealed the children's difficulties in comprehending relative clauses were artifacts of the experimental situation. Once the disturbing factors were removed, children displayed no difficulty in correctly understanding relative clauses." (p. 228).

The above examples of *wh*-questions and relative clauses, however, could all be taken as instances of movement of a *wh*-phrase past an intervening full (i.e. non *wh*) DP. The raised cases considered in Study 1 involved moving a full DP past another full DP. A cost function such as that in Gibson (2002), where cost of integration is defined only over the number of elements (DPs) crossed, predicts no difference in resources needed for a *wh*-phrase moving over a DP in an object-extracted *wh*-question versus movement of a full DP over a DP experiencer in a raised sentence. On the other hand, a source memory

[61] In the case of Sheldon (1974), both the grammatical function of the relativized DP inside the main clause (whether it was a subject or object), and the grammatical function of the relative gap inside the relative clause (whether a subject or object gap) were manipulated. Interestingly, while object-extracted relative clauses were more difficult for children than subject-extracted relatives when the relativized DP was in subject position, the opposite was found to be true for object DPs. When it was the object of the main clause that was relativized, object gaps in the relative clause were easier than subject gaps.

model such as that proposed by Gordon, Hendrik, and Levine (2002) would predict greater difficulty with the raised sentences than with *wh* sentences, since a *wh* word should interfere less with a full DP than would be the case for two full DPs. Could children simply be failing to comprehend sentences requiring extraction of a DP past a like DP?

This cannot be the case given at least two sets of acquisition findings. First, in V2 languages, children are found to topicalize objects to sentence initial position from very early ages. For example, Poeppel and Wexler (1993) find that German-speaking Andreas produces topicalized objects by age 2;1. Such sentences require movement of a full DP object past a full DP subject, yet such movement does not hinder the child. Second, local scrambling of a full DP object over a full DP subject in Japanese, a verb-final language, is also known to be early in acquisition (Otsu, 1994). In both cases of object topicalization and local scrambling, the movement of one full DP past another does not produce the same comprehension difficulties noted for raising a full DP subject past a full DP experiencer in Study 1.

The two cases noted above, however, involve *A'*-movement. One might wonder then if a more convoluted processing account could be told where a processing ceiling is reached only in the case of *A*-movement of a DP over an intervening element. Here the data and theoretical debate become much more complex and interesting. While certain currently available empirical data appear consistent with such a processing story, it is not at all clear why such an account should hold. Even if *A*-movement is more costly for the grammar than *A'*-movement, and note that the relevant *A*-chains are *shorter* than the *A'*-chains under discussion, why should a processing ceiling be reached with *A*-movement, but not *A'*-movement? How are the processing requirements of raising over an experiencer greater than for successive cyclic *A'*-movement in long-distance *wh*-questions, the latter which children comprehend? How is a processing account that hinges on *A*-movement over DPs being too hard for young children different from a grammatical account (i.e. UPR) that states basically the same thing? These are difficult questions to which we will return in later chapters. Meanwhile, what is now clear is that a simplistic processing account, by which children's difficulties with raising in Study 1 are attributed to reaching a processing ceiling when attempting to link syntactic elements across an intervening element in the surface representation, has been shown false on the basis of the data from Study 2. Children fail to comprehend raised sentences with either a sentence-initial or sentence-medial experiencer.

What then can be made of the particular pattern of child responses in Study 2, where performance on raised sentences with DR and ER foils is noticeably better than the chance level performance noted with MR foils? It would appear that children actually lack knowledge of raising, across all foil types, but are making use of an interpretive strategy that happens to allow them to select the correct picture with DR and ER foils even in light of their grammatical deficit, but does not work with MR foils, which leads to confusion and thus to chance guessing.

This notion that children apply a particular interpretation to a string that would otherwise be ungrammatical to them is nothing new. It is exactly what has been hypothesized in Section 2.2.4 when it comes to children's interpretation of raised sentences with a medial experiencer, namely, the *imagine*-analysis. The same general idea applies to children analyzing verbal passives (with non-subject experiencer verbs) as adjectival passives (Section 1.2.3). The central idea is that the child attempts a parse for any given input that is consistent with the current state of his grammar, where a particular syntactic analysis can stand in place for one that was unavailable due to lack of grammatical maturation. The child interpretation is considered a s(yntactic)-homophone for the adult analysis.

While the *imagine*-analysis captures the behavioral data from Study 1, one was left wondering why it applies. How do children come to such an analysis? Is fitting an interpretation to some input string something children are consciously thinking about during the experimental task? What about in natural settings when hearing the same sentences? What happens in (natural/elicited) production? These issues ultimately beg the question of what is an interpretive strategy. In the case of the *imagine*-analysis, one central question is to what extent interpretation may be based on children already having established a lexical entry for *seem* used in raised sentences, leaving aside the issue of the presence of the medial experiencer for a moment. If children truly believe that *seem* means something like (RtO) *imagine*, or even that all StS raising verbs have similar lexical semantics, then their performance on interpreting a raised sentence with a fronted experiencer will depend crucially on whether they accept a RtO verb with topicalized infinitival subject (predicts same pattern of performance as in Study 1) or whether they reject it (predicts chance performance across all foil types).

If children, however, do not have a preconceived notion of the meaning of *seem*, then perhaps they come to the experimental task with the flexibility to seek an interpretation consistent with the StS raised sentences used, whether they involve a sentence-medial or sentence-initial experiencer. Perhaps when the experiencer is in canonical medial position, the *imagine*-analysis applies, since it affords an interpretation consistent with both the elements in the input string (DP_1 V DP_2 $TP_{Nonfinite}$) and converges with an interpretation that allows consistent choice within the experiment (consistent, but ultimately incorrect, selection of the DR foils). The fact that this leads to consistent errors is of course an indicator that the interpretation is severely flawed, and probably accounts for why older children who have yet to acquire raising nonetheless appear to avoid it. When the experiencer is fronted, and if the *imagine*-analysis cannot apply, perhaps children will seek a different interpretation; again, one that is both consistent with the elements of the input string and maps to consistent choices in the experimental setting.

In one sense, this detailed discussion about what interpretation children are applying to StS raised sentences they clearly do not comprehend, at least in the adult sense, begs the question of whether it really matters. To a good extent it does, especially if our goal is to understand what children are doing in the face of an otherwise ungrammatical structure. Yet, if our primary goal is to simply discover cases of ungrammaticality, that is, to determine if StS raising for young children is grammatical or ungrammatical, then these questions of strategy certainly take a back seat. The results obtained so far *do* argue for the delayed acquisition of raising. That should be a central focus when discussing the details of these results. All said, however, it also seems that we would like to understand what interpretations are being employed to generate these particular results. Let us consider a few possibilities.

Start by noticing that children are not without *some* interpretive strategy when it comes to answering raised sentences with a fronted experiencer. While the chance performance results on raised sentences with MR foils indicate that children do not comprehend these sentences, children do not give up when attempting to interpret raised sentences with a fronted experiencer in all cases. If this were the case, children should respond at chance level to raised sentences with all three foil types. Such, however, is not what is observed. Rather, while indeed children are at chance level for MR foils, they demonstrate above-chance performance for DR and ER foils. That is, some interpretive strategy is available that allows children to successfully respond to raised sentences with a pre-posed experiencer-phrase in two of the three foil subconditions.

Could this strategy be the *imagine*-analysis that appears to account for children's responses to raised sentences with a medial experiencer in Study 1? As discussed above in Section 2.3.3, how the *imagine*-analysis is predicted to apply depends on whether children accept RtO sentences with topicalized infinitival subjects (41). If such sentences are fine for young children, then the *imagine*-analysis should be able to apply, in which case children should take the first (topicalized) DP to be the agent of the infinitival clause, and the second DP as the experiencer. Just as in Study 1, this interpretation should lead to children consistently choosing the DR foil (below-chance performance) since it matches in meaning, while leading to guessing (chance performance) with ER and MR foils, since neither these nor the correct picture would match in meaning. The experimental data from Study 2, though, clearly demonstrate that children do not manifest such a response pattern. DR foils are not comprehended at below-chance level, but are answered at above-chance level; ER foils are not comprehended at chance level, but at above-chance level. These findings are quite inconsistent with an *imagine*-analysis that assumes children can topicalize the infinitival subject with RtO verbs.

If, however, children do not accept RtO sentences with topicalized infinitival subjects (see Footnote 54 for comment), the *imagine*-analysis will not provide a parse that converges on a grammatical structure in English. As such, any child tied to the *imagine*-analysis would end up with a non-interpretable string, and should therefore demonstrate chance performance with all three foils, since random guessing would follow. Children, however, do not manifest this response pattern either. They are not at chance level for all three foils with raised sentences in Study 2, but rather just for MR foils. While children's responses in Study 1 are consistent with an *imagine*-analysis for raised sentences with a medial experiencer, their responses in Study 2 to raised sentences with a fronted experiencer demonstrate that something other than the *imagine*-analysis is being used. This means children are not tied to a lexical definition of *seem* meaning (RtO) *imagine*. Children do not know StS raising with an experiencer, but are trying various grammatical interpretations to attempt a parse depending on the nature of the particular sentence structure. With a medial experiencer, they arrive at something like the *imagine*-analysis. With a fronted experiencer, they are attempting some other analysis.

If premature children are not tied strictly to a particular lexical definition for *seem*, then perhaps they are free to search the input string for subparts that do fit frames with which they are familiar, and thus non-chance behavior might be elicited. For example, one can easily imagine that a child hearing the sentence in (48), with a matrix verb unknown to him, might still be able to resolve the CP, without comprehending how it fits with the matrix clause. Even parsing part of the sentence might lead the child to surmise that it involves Mary dancing. Not comprehending every part of a sentence does not entail that a child cannot extract some meaning from it.

(48) John *UNKOWN-VERB* that Mary is dancing.

A child hearing a sentence like (49) below, not comprehending raising, and not being able to apply the *imagine*-analysis (either because the child does not associate the verb *seem* with such an analysis, or because he believes it cannot apply with a fronted experiencer), might ignore the sentence-initial experiencer altogether, and still be able to focus in on the material after the pre-posed experiencer-phrase, in this case (50).

(49) To Mary, John seems to be dancing.

(50) John seems to be dancing.

Do we have any evidence to suggest that children might be ignoring the fronted experiencer with raised sentences? Indeed we do. We know that the children investigated in Study 2 are completely ignoring it. Or, at the very least, they are completely incapable of interpreting it as determining which character is the experiencer. If children could interpret it as such, then they should respond correctly for raised sentences with DR and MR foils, and be at chance with ER foils. Yet, this is not the observed pattern of results. Rather, children respond correctly to DR and ER foils, while answering at chance with MR foils. [Of course, children had no problems interpreting the fronted experiencer with unraised sentences.]

Given that the experiencer is being ignored in raised sentences, could it be that children are left to parse the remainder of the raised sentence, and are successful, such that StS raising is grammatical? That is, upon hearing (51), do children arrive at (52)?

(51) To Mary, John seems to be dancing.

(52) John seems to be dancing.

Such a strategy, where the child ignores the fronted experiencer-phrase in a raised sentence, for reasons yet unexplained, and then grammatically interprets the remaining portion of the raised sentence, actually would capture the findings for Study 2. Simply parsing (52) above should lead children to select the correct picture over either the DR or ER foils, since both foils contain the wrong character doing the action denoted by the embedded clause, and should lead them to chance performance for the MR foils, since both the correct picture and MR foil associate the subject with the action of the embedded clause. Yet, if it is correct that raising is grammatical, once the experiencer is ignored, one is in a real bind to explain the results from Study 1. Following such logic, should children not ignore the experiencer in a raised sentence like (53) and arrive at (54)?

(53) John seems to Mary to be dancing.

(54) John seems to be dancing.

Children, however, are not ignoring the experiencer-phrase in Study 1. Doing so would, as just mentioned, lead to chance performance on MR trials, and above-chance responses on DR and ER foils. Yet children are found to be at chance for ER and MR foils, and below chance for DR foils with raised sentences in Study 1. Therefore, an analysis that suggests children are ignoring the experiencer, and grammatically interpreting the raised sentence that remains, fails to unify the results of Study 1 and Study 2. Ultimately, if children *did* find the raised sentence that remained after ignoring the experiencer to be grammatical (e.g. as predicted by UFH), then they would be expected to succeed in StS raising studies that avoid the experiencer altogether (see Chapter 3).

Alternatively, children might make use of some strategy for interpreting an otherwise ungrammatical raised sentence once the experiencer is ignored. Since according to UPR (but not UFH), the child is hypothesized to not comprehend StS raising even when no experiencer is present, he cannot grammatically resolve the (raising) clause that remains after ignoring the experiencer-phrase, at least not in accordance with the adult representation. Yet, for a sentence like (54) above, a child might nonetheless associate the subject *John* with some event of dancing. This either could be through a non-grammatical direct association of the subject and predicate (e.g. some form of non-linguistic semantic linking), or by substituting the copula for the raising verb, resulting in the parse in (55).

(55) John is dancing.

This "copula-analysis" basically posits that children interpret the copula in place of a StS raising verb.[62] Note that the copula does not license an experiencer-phrase, neither in sentence-initial nor sentence-medial position (56). Therefore, the copula-analysis would seem likely only in cases where an experiencer is either absent or ignored. As discussed above, it is known that the children of Study 2 are indeed ignoring the experiencer in raised sentences.

(56) (*To Mary), John is (*to Mary) dancing.

If children do assume the copula-analysis for interpreting otherwise ungrammatical raising structures in Study 2, what results would be expected with respect to the various foils? With DR and ER foils, children should respond correctly, rejecting the foils since they do not have the subject performing the action of the embedded clause. In the case of the MR foils, where both the correct picture and the MR picture have the subject doing what is denoted by the embedded clause, children should be left to guess randomly, as neither picture is more "correct" than the other. These predictions are of

[62] The exact syntactic details of a grammatical version of the copula-analysis (as opposed to a non-syntactic linking of the subject and embedded predicate) are here purposefully left vague. While a few possibilities exist, there is little relevant evidence to help decide between them. The child could ignore both the matrix verb *seem* and the non-finite copula in the embedded clause, replacing the entire sequence with the finite copula (i). Alternatively, the child might only ignore the raising verb *seem*, leaving just the non-finite copula, which itself would then become tensed (ii). As discussed in Section 2.2.4, very young children know the difference between finite and non-finite clauses. Indeed, respecting finiteness distinctions was part of the original motivation for positing the *imagine*-analysis over a direct *think*-as-*seem* account. Given an otherwise ungrammatical string for the child, however, this reanalysis might not be so striking. Finally, the child could replace *seem* alone with the copula (presumably tensed, as was *seem*), resulting in a parse with two copulas, at which point one is ignored (iii).

(i) John ~~seems to be~~ dancing → John is dancing
(ii) John ~~seems~~ to be dancing → John ~~to be~~ dancing → John is dancing
(iii) John ~~seems~~ to be dancing → John ~~is to be~~ dancing → John is dancing

In any case, the critical issue for the copula-analysis is that it links the subject as the agent of the action denoted by the embedded predicate.

course exactly how children actually respond for raised sentences with a fronted experiencer-phrase. An analysis on which children treat the raising verb as the copula, and thus ignore the experiencer, would account for children's performance with raised sentences in Study 2. This analysis, though, also fails to unify the results from Study 1 and Study 2 by means of a single interpretive strategy that children are taken to use. Of course, there exists the very real possibility that children *are* using different analyses in both studies. Reasons for which children might not use the *imagine*-analysis with raised sentences with a fronted-experiencer have already been explored. An explanation for why children do not make use of the copula-analysis in Study 1 is therefore needed. One is readily apparent: interpreting the raising verb in (57) as the copula leads to the ungrammatical form in (58). Trying to use the copula-analysis with raised sentences with a medial experiencer does not provide an s-homophone.

(57) John seems to Mary to be dancing.

(58) *John is to Mary (to be) dancing.[63]

When it comes to children's interpretive strategies, a crucial criterion favoring one analysis over another would seem to be whether the particular strategy results in a parse that is "actionable". The parse should lead the child to an interpretation that maps on to some meaning that is operationally useful, that is, some meaning that allows for choice or differentiation in the physical world. With respect to the experimental studies, this mapping would be one that allows non-chance performance in the picture selection

[63] See Footnote 62 for comments about the grammatical implementation of the copula-analysis. Note that in whatever manner the raising verb and non-finite copula are dealt with in (54) such that they come to be replaced with a copula, the resulting sentence will not be grammatical with the medial experiencer.

tasks, at least for some foils. Children's grammar appears to want to avoid complete failure. It wants to avoid cases where guessing is always required due to interpretations that fail to converge on semantics that offer a meaning consistent with the world.

The *imagine*-analysis from Study 1 is actionable. The parse it provides leads to non-chance performance in the case of DR foils, to which it directly maps, leading children to consistently choose the wrong picture. That said, for two-thirds of the raised sentences it leaves children to randomly guess. There is no claim that children's grammar will always result in operational interpretations, only that grammar attempts to map between phonetic forms and the real world. The copula-analysis is also actionable. Once the experiencer is ignored, simply attending to the agent of the embedded clause allows non-random picture selection for two-thirds of the foils (ER and DR) in Study 2.

It is worth noting, however, that the mapping between the interpretations provided by the strategies and the pictures themselves is rather different. In the case of the RtO *imagine*-analysis, the putative interpretation is one that directly and exactly matches one of the two pictures (with DR foils). The copula-analysis paired with ignoring the experiencer, on the other hand, does not lead to such an exact match between interpretation and picture choice. The copula-analysis basically results in children choosing the picture (between correct and foil) that contains the non-experiencer DP engaging in the action mentioned in the embedded predicate. While this leads to success with two of three foil types, this analysis completely ignores the entire thought bubble depiction. The copula-analysis does not reference the experiencer at all. Children selecting between pictures are thus doing so while presumably ignoring a great deal of the pictures themselves (i.e. the "thinking" relationship). Crucially, however, neither strategy involves an interpretation that is inconsistent with both pictures (at least across all foils). While the copula-analysis might be seen to underdetermine which picture to choose, it is not incorrectly determining that choice in a consistent manner.

Before turning to the conclusions that can be drawn from Study 2, one further possible strategy for dealing with raised sentences and a fronted experiencer deserves considerable comment, as at first blush it might seem to offer a means for unifying the results of both studies. Take a raised sentence from Study 1 (59) and its interpretation as given by the *imagine*-analysis (60). Instead of *imagine*, however, consider a RtO verb like *expect*. If *seem* were being replaced with *expect* (61) instead of *imagine*, the results from Study 1 are still captured. Both *imagine* and *expect* have an experiencer matrix subject and infinitival subject that is the agent of the embedded clause. Children should therefore respond similarly to raised sentences in Study 1 whether they take *seem* to be *imagine* or *expect*.[64]

(59) Bart seems to Lisa to be waiving a flag.

(60) Bart imagines Lisa to be waiving a flag.

(61) Bart expects Lisa to be waiving a flag.

Unlike *imagine* (62), however, *expect* is grammatical without a second, post-verbal DP (63). In (63), *expect* is not a RtO verb, but rather, a subject control verb with the syntactic structure in (64). The interpretation here is reflexive, with the sentence in (63) meaning (65).

[64] The verb *imagine* is possibly a slightly better match than *expect* for the thought bubble depictions used in Study 1, but *expect* seems quite compatible as well.

(62) *Bart imagines to be waiving a flag.

(63) Bart expects to be waiving a flag.

(64) $Bart_i$ expects [PRO_i to be waiving a flag].

(65) Bart expects himself to be waiving a flag.

From prior discussion, it was noted that children are ignoring the experiencer (as such) in Study 2 where it is fronted to sentence-initial position. If children *were* (minimally) interpreting the experiencer-phrase as denoting which character was doing the thinking in the test pictures, they should be above chance for DR and MR foils, but at chance for ER foils. This response pattern, however, is not what is found. Instead, children are above chance for DR and ER foils, but at chance for MR foils.

The issue then becomes one of determining how children interpret the linguistic material that follows the initial experiencer-phrase. Ignoring the experiencer in (66) leaves (67). The copula-analysis, whereby children analyze (67) as (68), does appear to capture the experimental findings, but at the expense of failing to achieve a unified account with the *imagine*-analysis, which does not apply to (66) (see previous discussion) and cannot apply to (67), as seen in the ungrammatical (69).

(66) To Lisa, Bart seems to be waiving a flag.

(67) Bart seems to be waiving a flag.

(68) Bart is waiving a flag.

(69) *Bart imagines to be waiving a flag.

Returning to the case of *expect*, unlike with *imagine*, it can be substituted in (67) to produce the very grammatical subject control sentence in (70).

(70) Bart expects to be waiving a flag.

Since RtO *expect* in place of StS raising *seem* captures the data in Study 1, it is worth seriously considering if subject control *expect* could likewise explain the data in Study 2. This "*expect*-analysis" is a mixed approach based around a single lexical item, having different syntactic analyses depending on use. As such, it might serve to explain both sets of acquisition data. Of course, it must first be asked whether the *expect*-analysis can capture the data from Study 2.

First, the *expect*-analysis provides a reflexive parse for the raised sentences in Study 2 once the experiencer is ignored. That is, the logical matching picture for (70) would be something like that in Figure 2.3.1. Obviously, no foils were included that match just such an interpretation (since the *expect*-analysis was not considered or anticipated prior to beginning experimental testing). Thus, no experimental control was used that could directly test for the *expect*-analysis. Unlike the copula-analysis, which was noted to underdetermine which picture to select (i.e. the resulting interpretation contains less information than depicted in either picture), the *expect*-analysis would actually lead to interpretations that no pictures (correct or foil) match.

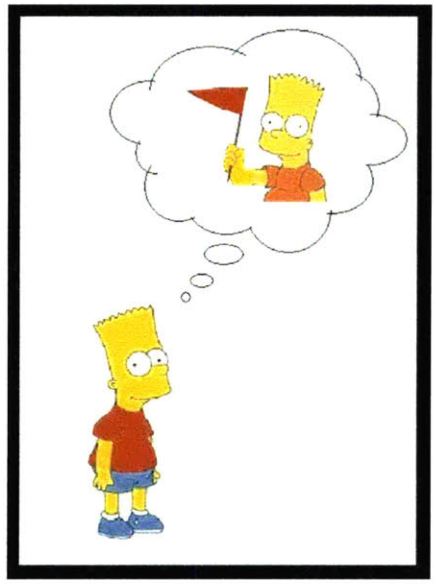


Figure 2.3.1: Example of a possible reflexive foil (untested) for the raised sentence *To Lisa, Bart seems to be waving a flag.*

Perhaps then *if* reflexive pictures had been included, children would have consistently chosen them. Since they were not, and if children nonetheless stuck to an *expect*-analysis, it could be that they still try to choose the picture that is most consistent with their linguistic interpretation. This requires a careful examination of each picture to determine which "best" matches the derived semantic meaning of the test sentence.

The *expect*-analysis makes two predictions: one about which character is the experiencer (doing the expecting/thinking) and one about which character is the subject of the embedded clause (doing the action of the embedded predicate); under this analysis, the same character is doing both. Given a test sentence in Study 2 like (66) above, the expect-analysis results in a parse like (70). Since neither the correct picture for (66) nor any of the three foils match the (reflexive) meaning of (70), children might thus focus on just one of the two predictions derived from the *expect*-analysis. If children in Study 2 are subject to the *expect*-analysis and focus only on which character is doing the thinking (the character with the thought bubble), then they should consistently respond incorrectly to DR and MR foils, since the *expect*-analysis picks out the wrong experiencer. Such children would be at chance for ER foils, since neither the correct picture nor the ER foil depict the experiencer given by the *expect*-analysis. If instead, however, children subject to the *expect*-analysis focus only on which character is the agent of the embedded clause, they should consistently respond correctly to DR and ER foils since this analysis does correctly determine the agent, and should respond at chance for MR foils since both the correct picture and MR foil depict the same (correct) character performing the action denoted by the embedded predicate.

As already noted for Study 2, prior to acquiring StS raising (i.e. not above chance for raised sentences with all three foils), children are above chance for DR and ER foils, but at chance level for MR foils. The prediction of the *expect*-analysis (with the added caveat that children pay attention to the agent, perhaps since no picture faithfully matches the reflexive meaning) matches very closely the data gathered in Study 2 from the four- to six-year-old children on raised sentences with a fronted experiencer (Table 2.3.8).[65]

Of course, one must ask why children would only pay attention to the agent given the *expect*-analysis (as opposed to the experiencer, which would result in completely different performance than what was found experimentally). It is not at all obvious that one should track only the agent given an *expect*-analysis parse and the pictures employed in Study 2. To investigate this issue, a short experiment was conducted with 45 adults (to match the 45 four- to six-year olds from Study 2 who might be subject to the *expect*-

[65] Since most (12/15) of the seven-year-old children in Study 2 have acquired raising (above chance on all three foil types), they have been omitted from Table 2.3.8 since they would not bear on the question of whether or not the *expect*-analysis applies when StS raising has yet to be acquired.

analysis). Each adult was sent a (.pdf) file containing three sets of test sentences and pictures (see Appendices A3). The sets consisted of one test sentence and two pictures, and adults had to choose which picture best matched the sentence. The actual sentences and pictures were taken from Study 2, with modification. At random, among the items with raised sentences, one DR, one MR, and one ER item were selected. The raised sentences were then modified to match the *expect*-analysis: the fronted experiencer-phrase was deleted, and the verb *seem* was replaced with *expect*. The test sentences are therefore grammatical sentences in English, but of course, match neither picture with which they are paired, as also would be the case for children subject to the *expect*-analysis (since neither picture is reflexive in the sense of Figure 2.3.1). Adults were instructed that they nevertheless had to choose one of the two pictures as matching best.

The question being addressed with this small, informal test is whether or not adults given explicit *expect* sentences (namely exactly those predicted to follow from the *expect*-analysis) would respond in a manner consistent with only focusing on the agent of the embedded clause, and thus consistent with children's actual behavior in Study 2. Adult performance is noted in Table 2.3.8, where adult responses were taken to be "correct" if the picture chosen matched the one that should have been chosen if the original raised sentence in Study 2 had instead been given (i.e. the sentence that the children in fact were given). As can be seen, adult performance on the ER item is very similar to children's performance with such foils (88% vs. 84%), as is also the case for the MR foils (53% vs. 60%).

Foil Type	E-A: Thinker	E-A: Agent	Study 2	E-A: Adults
R-DR	0%	100%	86%	38%

R-ER	50%	100%	88%	84%
R-MR	0%	50%	60%	53%

Table 2.3.8: Predicted patterns of performance for the *expect*-analysis (E-A) if children only pay attention to the experiencer ("thinker") or the subject of the embedded clause ("agent"), compared to the performance on raised sentences of the four- to six-year-old children in Study 2, and to the performance of an adult group given actual *expect* sentences.

That is, adults are consistently choosing one of the two pictures in the ER case, but are at chance in the MR case (remember, neither picture matches the test sentence in either case). The picture chosen in the ER case depicts the named character engaged in the action of the embedded clause, while both pictures fail to show the named character doing the thinking. For the MR case, both pictures have the named character as the agent of the embedded predicate. At first blush, then, perhaps it is the case that given an *expect* sentence and non-reflexive pictures, subjects (both adults and children) choose based on the agent. Unfortunately, adults are at chance level for the DR case.[66] In the DR case, one picture has the named character performing the action while in the thought bubble, wheras the other picture has the named character doing the thinking while the other character performs the action. If responses were made on the basis of agentivity, adults should consistently choose the picture with the named character as agent (which happens to be the correct picture for the raised sentence which served as input to the *expect*-analysis). Instead, the trend for adults (though not statistically significant) is to actually choose the picture in which the named character is doing the thinking in the DR case.

[66] The 38% is just barely at chance (as opposed to below-chance) level, where 37.7% is the cut-off given a 95% one-tailed confidence interval.

It is unclear what selection criteria adults must be employing to derive this pattern of responses across all three test sets, but crucially, a direct test of the *expect*-analysis (in adults) fails to demonstrate a natural preference for consistently selecting based on agentivity as required if the *expect*-analysis is to explain the findings of Study 2 (and Study 1). While this quick test cannot be taken as conclusive (only three items tested, no randomization of item presentation, lack of consistent experimental setting, etc), it does suggest that an agent-centric *expect*-analysis is certainly not predetermined.

Altogether, the data point more straightforwardly to the copula-analysis than the *expect*-analysis for raised sentences with a fronted experiencer. While the *expect*-analysis would integrate the findings from both studies under a single theorem, it requires several caveats (most crucially the agent-based selection criterion which was not confirmed in an adult test group). The copula-analysis, on the other hand, does not require any further interpretive juggling. If, however, children are making use of the *imagine*-analysis in Study 1 and the copula-analysis in Study 2, it suggests a certain cognitive and linguistic flexibility on the children's part, such that they are able to seek different interpretations based on the particular syntactic details of the test sentences. This would mean that children are not coming to the experiments with predetermined meanings of raising verbs (like *seem*), but are extending interpretations based on a combination of known syntactic frames and preconceived sense of partial meaning (e.g. that *seem* has something to do with *thinking* or appearance). On the other hand, none of the data can rule out the possibility that indeed children's lexical item *seem* is subject to the *expect*-analysis, but in Study 2 this analysis fails to map onto any pictures, and as a result, only then do children seek a copula-analysis. Finally, there might simply be a mix of children, some of whom employ the *imagine*-analysis in Study 1 and the copula-analysis in Study 2, while other children employ the *expect*-analysis in both cases (which might either lead to paying attention only to the agent, or seeking out the copula-analysis, depending on how they deal with the lack of reflexive pictures in Study 2).

Ultimately, what are needed are tests with the experiencer-phrase explicitly missing. These experiments should furthermore directly investigate both the copula-analysis and the *expect*-analysis. For the moment, however, the details of strategic interpretation will be placed on the back burner, while the issue of why a raising deficit might exist is examined. In particular, attention will be turned to whether the grammatical issue underlying children's difficulties and delays with verbal passives also lies at the heart of the observed problems with raising. While a general similarity in age of acquisition hints at a common underlying explanation, one can do better with within-child comparisons of acquisition patterns for both structures.

2.3.8 *Conclusion*

Results from Study 2 demonstrate that even with a fronted experiencer, children continue to manifest difficulties in comprehending raised sentences. An examination of the developmental curves for raising with a fronted experiencer match those for raising with a medial experiencer. Furthermore, successful comprehension of the raised sentences in Study 2 is reached for a majority of children in an age group only with the seven-year-olds, which is also when raising over an experiencer-phrase was acquired in Study 1. Notably, this is about when some researchers have argued verbal passives are acquired.

As for the errors that children make in attempting to interpret raised sentences with a fronted experiencer, it seems that they are not applying the *imagine*-analysis from Study 1. Instead, the data are consistent with children making use of a copula-analysis once the fronted experiencer is ignored. If so, this suggests that children are not decisively tied to a particular lexical definition for *seem*, as used in StS raised sentences, such that different interpretive strategies would otherwise not be available. Children would thus appear to remain flexible in their interpretation of what otherwise would be ungrammatical raised sentences. Alternatively, the results from both studies could be captured under the single *expect*-analysis, though it remains unclear if it is applicable when a reflexive interpretation is reached in the case when no post-verbal DP is present.

These data serve to rule out a processing account working over surface representations. It is not the case that raising is delayed only when a dependency is established in the surface representation that crosses over an intervening DP. The question thus becomes: what accounts for children's delayed knowledge of raising as observed in Studies 1 and 2? Two grammatical accounts, UPR and UFH, both predict this delay. One of these accounts, UFH, ties comprehension difficulties concerning subject-to-subject raising to the presence of the experiencer-phrase. According to UFH, the syntactic operation of raising over an experiencer is ungrammatical for premature children, but not StS raising itself; whether that experiencer is in medial or initial position is irrelevant. UPR, though, clearly predicts ungrammaticality for StS raising regardless of the presence, position, or absence of the experiencer.

There are two experimental approaches that one can follow in order to further examine the validity of the grammatical accounts: generally, as opposed to non-grammatical approaches to accounting for delayed comprehension of raising, and specifically, investigating the validity of different grammatical approaches. A direct approach to investigating these issues is to simply eliminate the experiencer-phrase completely, and see how comprehension of raising is affected, and this path will be pursued in Chapter 3. An indirect approach comes from asking whether the course of acquisition for StS raising mirrors that for verbal passives. Both UPR and UFH predict similar developmental trajectories for these two structures (for UFH, only when the experiencer-phrase is present), as well as within-subject correlations for knowledge of each. It is to this latter, indirect approach that we turn first.

2.4 Study 3: Relationship Between Raising (over an Experiencer) and Verbal Passives

2.4.1 *Motivations*

Having found that children are delayed in the acquisition of StS raising, at least with an experiencer-phrase present (whether in canonical position or fronted), we now turn to the question of whether such delay is due to the same underlying cause as that giving rise to children's delayed acquisition of verbal passives. Studies 1 and 2 both find a jump in comprehension of StS raised sentences between six and seven years of age. This matches the window for passive acquisition noted in previous studies (e.g. Maratsos, Fox, Becker, and Chalkey, 1985; Hirsch and Wexler, 2006a, Hirsch and Wexler, 2006c; see Section 1.2.2 for a review). While the general age of acquisition for both structures appears to be roughly similar, it is unknown whether their development is tied at a more fundamental and explanatory level. To be sure, maturation theories such as UPR and UFH both predict, for reasons to be reviewed below, that verbal passives and StS raising (with an experiencer only for UFH) should be linked in development. It could be, however, that we are merely witnessing two independent grammatical issues in development which just happen to roughly correspond with respect to age of acquisition. Whether the actual developmental curves for both structures match across ages, and whether the development of both structures is the same within individual children, cannot be addressed on the basis of the current empirical data.

UPR predicts that verbal passives and raising will both be delayed due to a ban on defective (i.e. weak) phases in premature child grammar. According to Chomsky (2001), passive and raising *v*Ps are both defective in adult grammar. For passives, if the *v*P were non-defective, this would entail that the base-generated object DP could not raise to subject position. In the case of StS raising, the embedded subject could not raise to matrix subject position. UPR posits that premature child grammar does not allow non-defective phases; all phases are strong phases. As such, passives and raising structures are both ungrammatical for premature children under UPR.

These two syntactic structures are also ungrammatical for premature children according to UFH (in the case of StS raising, only when the experiencer is present), but for different reasons than those suggested by UPR. UFH disallows movement from already moved constituents in premature child grammar, which renders Smuggling ungrammatical (Collins, 2005a, 2005b). Passives are assumed to involve Smuggling of the object past the external argument *by*-phrase. Raising over an experiencer also requires Smuggling in order to get the embedded subject past the experiencer. Collins' analysis of raising does not require Smuggling when no experiencer is present, so that UFH does not predict delays for raised sentences without an experiencer, a prediction that will be taken up in Chapter 3. For the moment, we deal only with the predictions as applied to raising over an experiencer, for which we now have actual acquisition data.

Thus, UPR and UFH predict that both StS raising (over an experiencer) and verbal passives will be delayed for children whose grammar has yet to mature to the adult form. These two grammatical structures, therefore, should be delayed to the same extent in child development. This means more than being able to observe similar developmental curves for both structures. Rather, these two theories strongly predict that acquisition of each structure should correlate not just among groups of children, but crucially, within individual children; a child's grammar, *ceteris paribus,* should either rule both structures grammatical or ungrammatical. That is, at a given point in time, a perfect correlation reflecting grammaticality should obtain within individual children.[67]

If delays in raising and passive acquisition are due to unrelated grammatical development or differences in learning, then there is no reason to expect such strong within-subject correlations to exist. If raising and passives are subject to idiosyncratic developmental pressures, even if those pressures roughly relieve around the same age, one would expect to find children who have acquired one structure but not the other. If, for example, the raising over an experiencer results found in Studies 1 and 2 are due to a processing, as opposed to grammatical, difficulty, then one would likely not expect

---

[67] Barring of course any necessary learning that must accompany syntactic development as would be relevant for successful comprehension. In the case of StS raising this would, for example, mean having learned to correctly categorize a StS raising verb as such, and for passives, for example, having learned the correct morphology indicative of passive syntax. This type of learning must take place before children can demonstrate knowledge of the relevant syntax, even if the potential to comprehend the syntax has already matured.

comprehension of passive and raised sentences to correspond, since the processing issues related to passives and raising do not mirror each other, at least not to the obvious exclusion of other structures known to pose no problems to children of the relevant ages. Also, if the raising comprehension difficulties noted in Studies 1 and 2 were simply due to some association with the presence of the experiencer-phrase, then it is unclear how such an issue could be related to problems with passives, which do not contain any such experiencer-phrase.[68]

Overlapping developmental curves for both raising and passives are predicted by UPR and UFH. The time courses of development for both structures should mirror each other over time, once extraneous factors, such as compensatory strategies, are taken into account. Moreover, these maturation theories predict comprehension for raising and passives to be evident within individual subjects. Thus, within-subject testing is required to validate such theories. As such, it will be necessary to administer both raising and passive comprehension experiments within a very short time frame in order to be able to investigate these predictions for within-child grammatical correlations.

2.4.2 *Experimental Design*

In order to investigate the relationship between the acquisition of StS raising and the acquisition of verbal passives, the same 70 children who participated in Study 1 were also administered a test of verbal passive comprehension within two weeks of having

[68] Note, as described below, that UFH does predict that raising comprehension should correspond to passive comprehension for children, but that only in the case of raising over an experiencer. This is very different than a (so far strawman) theory that states raising in Studies 1 and 2 was delayed only because of the experiencer (for reasons left vague), such that raising without an experiencer should be fine. Such an account makes no prediction that raising and passive comprehension should correlate, whether the experiencer-phrase is present or absent in cases of raising. UFH makes strong predictions concerning such correlations, both when the experiencer is included (strong correlation should exist) and when the experiencer is omitted (no correlation should exist).

participated in Study 1.[69] Study 3 tests comprehension on four conditions, crossing voice (active vs. passive) and verb type (actional vs. subject experiencer (SE)), to produce the four sentence types in (71)-(74).[70]

(71) Bart pushes Lisa.

(72) Lisa loves Homer.

(73) Marge is pushed (by Bart).

(74) Homer is loved (by Marge).

Eight verbs were used, consisting of four actional verbs (*push, kiss, kick, hold*) and four subject experiencer verbs (*remember, love, hate, see*). Eight items were constructed for each active condition, and 16 items for each passive condition. All sentences were semantically reversible, meaning each sentence was such that either character could logically have performed the action denoted by the verb (c.f. *The man ate the apple*, which is not semantically reversible, since apples do not typically eat men). In order to

[69] The administered passive test is that developed by Hirsch and Wexler (2006a), which has also been used successfully, with slight modification, obtaining similar results by Perovic and Wexler (2007), Hirsch, Modyanova, and Wexler (2006), and Crawford and Hirsch (2008).

[70] The passive conditions were further subdivided according to whether or not they contained a *by*-phrase, half of the passives being full passives (with a *by*-phrase) and half being truncated passives (without a *by*-phrase). It has been claimed by Fox and Grodzinsky (1998) that children comprehend truncated subject experiencer passives, which they take as evidence that children do not have a general deficit in passive comprehension. Their claim that truncated subject experiencer passives are not delayed, based on only eight children, has since been shown to be false. Not only has their finding not been replicated, but numerous studies using both more children and more items demonstrate that truncated subject experiencer passives are just as delayed as their full counterparts (Gordon and Chafetz, 1990; Hirsch and Wexler, 2006a; Hirsch and Wexler, 2006c). Of the 140 subjects examined by Hirsch and Wexler (2007c), only two children have scores matching the Fox and Grodzinsky pattern of good performance on truncated subject experiencer passives and poor performance on full subject experiencer passives. Furthermore, the particular affector *by*-phrase account proposed by Fox and Grodzinsky has also been refuted (Hirsch and Wexler, 2006b). Fox and Grodzinsky's result appears to be due to methodological problems, including too few subjects, too few experimental items, a failure to randomize conditions, and delays between when conditions were tested.

minimize task demands, once again only four *Simpsons* cartoon characters were used throughout the experiment, with which the children were familiarized in the introduction to the study.

To assess children's comprehension of these four sentence types, we employed a two-choice sentence-picture matching task wherein children were shown on a laptop screen two pictures side-by-side depicting opposite events. So for example, for sentence (71), a child would have been presented with two pictures, one showing Bart pushing Lisa, the other showing Lisa pushing Bart (Figure 2.4.1).[71] Children were told to choose the picture best matching the sentence they were read, after which their answers were logged on the computer by the experimenter before continuing to the next item. All sentences were read twice to the child before he was allowed to respond. The location of the correct picture (left or right side of the screen) was balanced across the individual verbs, conditions, and the entire experiment. Items were presented in a randomized order to each child. See Appendices (A4) for details.

[71] Three of the four subject experiencer verbs were depicted using thought bubbles (all except *see*). As mentioned in Section 2.2.1, previous research has demonstrated that children comprehend such pictorial representations at the ages under investigation here. Furthermore, numerous studies, listed in Footnote 69, have successfully used these same pictures to investigate the acquisition of verbal passives, finding children to have no difficulties with subject experiencer verbs in the active voice. Thought bubbles were also used in Study 1, causing no issues for nearly all children.

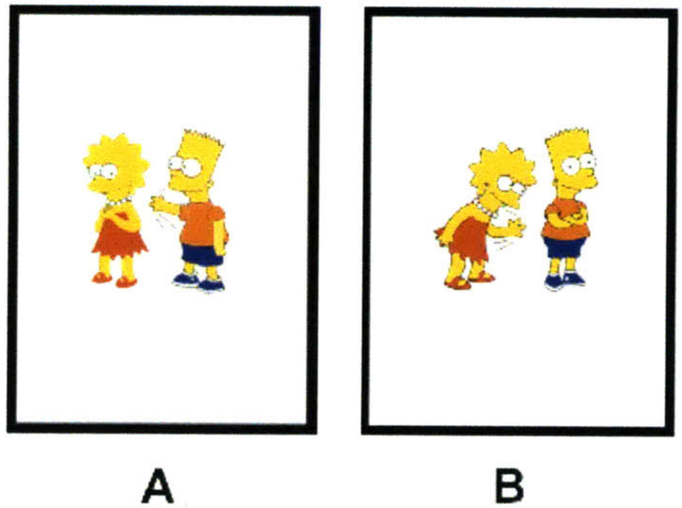


Figure 2.4.1: Example pictures in Study 3. Children had to choose between the picture on the left (A) and the picture on the right (B) as to which best matched the test sentence read aloud by the experimenter.

2.4.3 *Results*

The results of this passive experiment appear below in Table 2.4.1. Children performed extremely well on both active conditions. Excellent performance on the subject experiencer actives, which were actually comprehended slightly better, but not significantly so, than their actional counterparts, indicates that children have no general problem comprehending subject experiencer verbs, nor any difficulty with the experimental methods for assessing such knowledge. We replicate all past experiments crossing voice and verb type, and find much worse performance for subject experiencer passives as compared to actional passives. Furthermore, the study replicates previous findings that subject experiencer passives are not comprehended until around seven years of age (see Section 1.2.2)

Age Group	Act. Actives	SE Actives	Act. Passives	SE Passives
3	96.3%	98.8%	65.6%	38.1%
4	95.0%	98.8%	86.3%	50.0%
5	97.5%	98.8%	92.5%	58.8%
6	97.5%	98.8%	89.4%	45.6%
7	97.5%	97.5%	95.6%	75.6%
8	100.0%	98.8%	92.5%	82.5%
9	100.0%	100.0%	95.0%	90.6%
Average	97.7%	98.8%	88.1%	63.0%

Table 2.4.1: Accuracy for all conditions across all age groups for the passive experiment portion of Study 3. [Act. = Actional, SE = Subject Experiencer]

For the purpose of examining in detail the relationship between the acquisition of raising and that of passives, it is vital to minimize the influence of any compensatory strategies that children might be employing in an attempt to comprehend structures for which their grammar is unable to provide a representation. The goal is to separate the study of core grammar from the use of strategies. Since these strategies are hypothesized to be dependent on environmental factors and mechanisms of general cognition, and not on pure syntactic competence, they must be isolated for the purpose of exploring deeper syntactic knowledge.[72] For raising, this means examining only the data from MR and ER trials since children employ a consistent (albeit wrong) strategy when given DR foils (i.e. the *imagine*-analysis). Including responses to DR foils would therefore bias comprehension scores *down* for children having not acquired StS raising. As for passives, only subject experiencer passives should be included, as actional passives are "understood" using the adjectival strategy, and including actional passives would therefore bias passive comprehension scores *up*. With respect to correlating children's

[72] The fact that children's analysis of actional passives is at least in part determined by environmental factors is reviewed in Section 1.X.

actual accuracy scores, it makes most sense to avoid such biased scores.[73] Thus, in all subsequent analyses, performance on raising will be examined only with respect to comprehension of raised sentences paired with MR and ER foils. Similarly, for passives, only scores on subject experiencer passives will be examined further.

Plotting children's accuracy scores on raising and passives by age group reveals very similar developmental curves for the two grammatical structures (Figure 2.4.3). Before seven years of age, children show generally no improvement in comprehending either raising or passives, with sudden and dramatic improvement at age seven across both conditions. It is also only with the seven-year-old group that there is a noticeable deviation from chance level (50% accuracy) for both curves.[74] The similarity of the two developmental curves is at least suggestive of a single grammatical deficit underlying delayed acquisition for both structures.

[73] When it comes to examining the relationship between above-chance comprehension of raising and passives, the inclusion of DR foils for raising would be acceptable (the "bias" is no longer relevant; it plays basically no role in determining which children are above or not above-chance comprehension levels). Actional passives would still need to be avoided, however, as they would continue to bias children's passive scores, suggesting that some children have acquired verbal passives, when in fact they have only acquired the adjectival strategy. For simplicity, we shall only include responses to raised sentences with ER and MR foils when examining both accuracy and above-chance performance throughout this discussion of Study 3.

[74] Note that this analysis and graph includes data from all children who participated in Study 1, not just those who succeeded with *think* and unraised items. Regardless of why raising was not understood by certain children (e.g. difficulties relating to issues assessed by performance on *think* and unraised conditions), these children do not comprehend StS raising. Thus, the developmental curves need not necessarily reflect performance on raising solely due to grammatical deficit or acquisition. That said, the striking similarity between the shapes of the two curves is worthy of serious consideration.

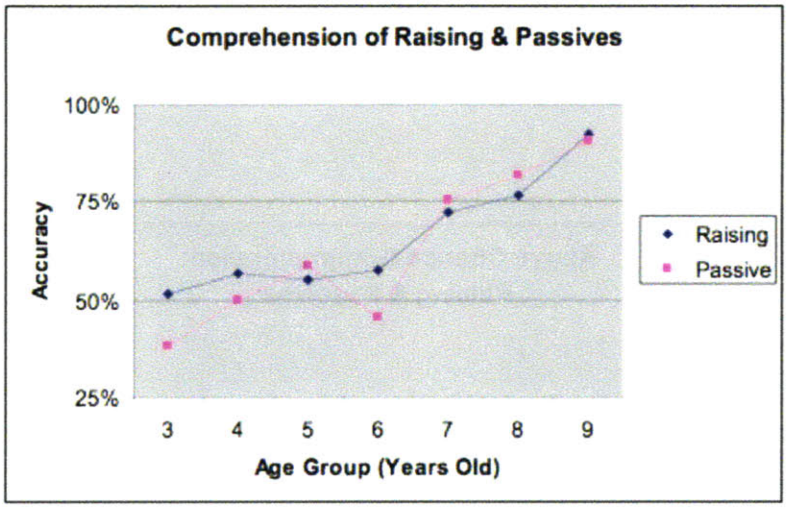


Figure 2.4.3: Similar developmental curves for comprehension of StS raising and verbal (subject experiencer) passives as measured by age group accuracy in Study 3.

The similarity in the patterns of acquisition is further demonstrated by considering the number of subjects in each age group who score above chance on each structure.[75,76] As reflected in Figure 2.4.4, before the age of seven, no more than three of ten children score above chance on either structure in any age group. The seven-year-old group, however, has six children (100% increase) who score above chance on raising and passives. This sudden increase in above-chance comprehension accounts for the sudden increase in group accuracy noted in Figure 2.4.3 above, as opposed to all children doing just slightly better. That is, increases in group accuracy across age groups is due to an increased number of children performing at above-chance level, not simply to all children within the age group doing marginally better on each condition. Once again, the strong

[75] Note that statistically significant above-chance performance here is different than in Table 2.2.3 (Section 2.2.3), as responses to DR items are not being considered now. Here, statistically significant above-chance performance for raising is at least 75% correct (9 of 12 items correct). For passives, statistically significant above-chance performance is defined as at least 75% correct on both truncated and full subject experiencer passives (6 of 8 items correct on both subconditions), which is relatively conservative.
[76] See Footnote 14.

similarity between these developmental curves hints at a common underlying cause for both structures being delayed in acquisition.

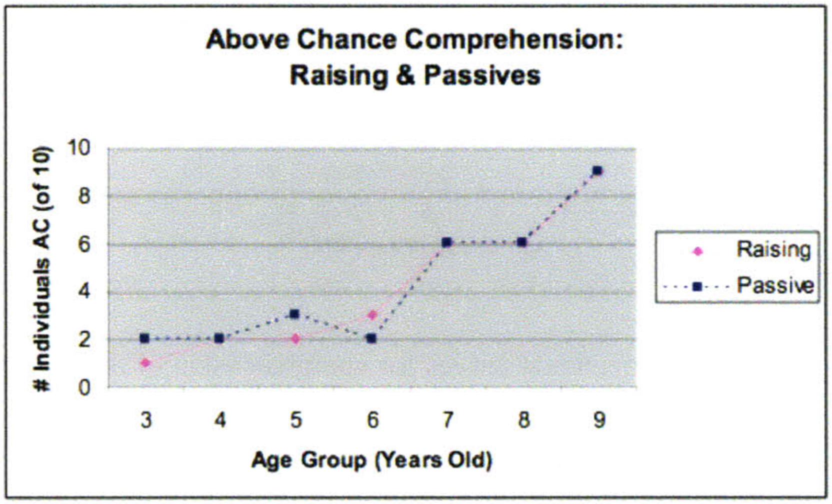


Figure 2.4.4: Similar developmental curves for comprehension of StS raising and verbal (subject experiencer) passives as measured by children in each age group scoring at above-chance level in Study 3.

While such developmental data strongly suggest that the acquisition of raising and passives are fundamentally linked, providing exciting evidence in support of certain maturation theories, what such theories predict is not simply that the average age of acquisition for both structures should match (in this case somewhere between six and seven years of age), but rather, that acquisition of both structures should correlate within individual children.

It is important to note, however, that the maturation theories do not predict that a strictly linear correlation should hold between children's actual scores on StS raising and verbal passives. That is, under the maturation theories, there is no reason to expect that if a child gets 66% of raised sentences correct that he should get exactly 66% of passive sentences correct (e.g. as opposed, say, to 44% correct). Rather, these theories predict that only two groups of children should be observed. The first group would consist of children who have not undergone the relevant maturation, and who should thus score poorly (i.e. not above chance) on both raising and passives. The second group of children would be made up of those whose grammar has matured, such that these children are expected to comprehend (i.e. score above chance for) both raising structures and passives. Before grammatical maturation takes place, comprehension of raising and passives is guided by relatively idiosyncratic and independent strategies (i.e. the *imagine*-analysis for raising and the adjectival analysis for actional passives, plus of course cases of random guessing). Thus, there is no great expectation of a strong correlation between the actual scores on raising and passives, especially for the younger premature children. Rather, the strong prediction of UPR and UFH is that there will not be children who comprehend raising (over an experiencer) but not passives, and vice versa. In other words, the maturation theories predict very strong correlations of above-chance performance on raising (over an experiencer) and verbal passives.

Before examining such correlations, it is necessary to remove from consideration those children who failed to comprehend the *think* and unraised *seem* trials during the raising experiment, as measured by performance on MR trials. Problems with Theory of Mind or simply in understanding (unraised) *seem* guarantee problems with raising, even if the relevant linguistic maturation has taken place to make passives and raising grammatical. Removing such cases leaves data from 55 children.[77]

[77] In Hirsch and Wexler (2007a), only 53 children were included in the subsequent correlational analyses. In error, two three-year-old children were accidentally omitted from their analyses. Both children did

The data for these remaining children is summarized in the scatter plot below (Figure 2.4.5), where lines at 75% indicate above-chance level for raising and passives.[78] Statistically significant and very high correlations for raising and passives obtain when either children's exact scores ($r(53) = 0.804$, $p < 0.0001$) or above-chance performance ($r(53) = 0.857$, $p < 0.0001$) is examined.[79]

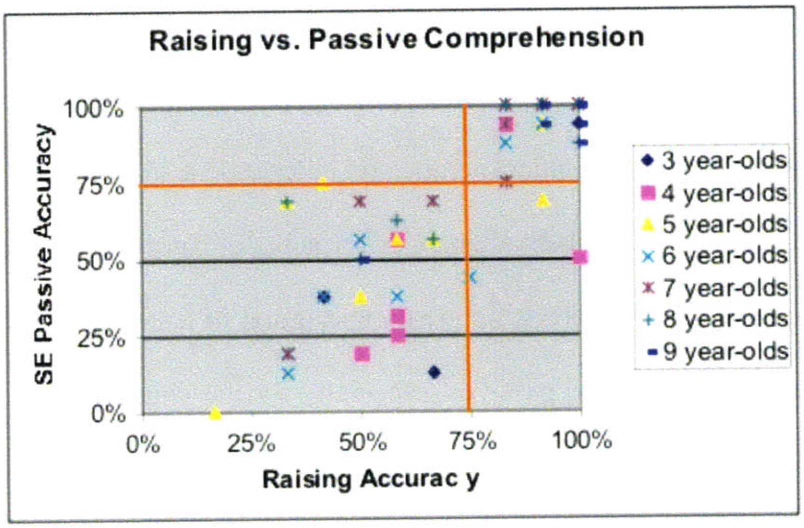


Figure 2.4.5: Scatter plot for comprehension scores from SE passives and StS raising for 55 children from Study 3. Notably, two groups of children are observed: those in the lower left quadrant who do not know either structure, and those in the upper right quadrant who know both structures. Orange lines at 75% indicate statistically significant

comprehend the relevant *think* and unraised sentences with MR foils, but their data was unfortunately ignored. Including this data serves to further the evidence for the maturational theories under consideration.

[78] See Footnote 14 for how above-chance levels are statistically derived.

[79] Calculating a Pearson (r) correlation in the case of the binary (above or not above chance) data is a bit suspect statistically. Such a parametric test is not really intended for binary data (though the test is rather robust). Another way of trying to show the same thing as the correlation is obtained by asking how children's performance *corresponds* between structures. This can be done by examining whether a match or mismatch is noted between each condition with respect to above or not above-chance performance for each child. Correspondence is defined as total cases of match divided by total number of children. In this case, correspondence between raising and passives stands at 92.7% (51/55).

above-chance level of comprehension for each structure. [Note: Many data points overlap.]

The scatter plot shows that in general, older children (seven years-old and older) tend to cluster in the upper right quadrant (above chance on both structures), while children younger than seven years tend to populate the lower left quadrant (not above chance for either structure). As predicted by UPR and UFH, very few children fall outside these two quadrants. There are only four apparent contradictions to the predictions of these maturation accounts. Upon further review, three of these turn out to be very marginal challenges. Had one subject missed just one more passive item he would not be an exception (not above chance for either structure). Had another subject scored one item better on passives, he would not be an exception (above chance for both structures). Had the third child scored one item worse on raising, he would also not be an exception (not above chance for either structure). That is, had these three children scored just one item differently they could not be considered exceptions. Given young children's already discussed penchant for distractibility, the relevance of a single item over the course of *two* rather extensive testing sessions is hardly a smoking gun in the face of otherwise very strong support for the correspondence of these two syntactic structures.

That leaves only one apparent "true" exception, a four-year-old child who got all of the StS raising items correct, but was at chance for the subject experiencer passives. While he could represent a real exception to what otherwise seems a perfect correlation between raising and passive scores, very plausible alternative explanations exist. Perhaps this child has undergone the relevant linguistic maturation, but for some unknown reason, fails to recognize the morphological markers of passives. That is, if the child had undergone the necessary maturation such that his grammar could represent passives, but nonetheless, had not learned the morphology signaling passive syntax (i.e. *be* plus past participle, and *by*), his behavior would be totally understandable, and even predicted. If this were true, he would not constitute an exception to the predictions of the maturational accounts.

The other possible explanation for this "inconsistency" also totally compatible with maturation accounts, is that the child was actually premature for both structures, but that random guessing enabled him to guess correctly for the raised sentences. This idea is quite natural given expectations for guessing and a random chance model. On the binomial distribution, the expected probability for scoring above chance on 16 trials (all the ER and MR foils), assuming chance guessing (50%), is 3.84%. This seems pretty remote, but remember, there are 29 children who did not score above chance on both passives and raising. Thus, for those 29 children, it is expected that one child (0.0384 * 29 = 1.11) will score above chance on raising even though he is actually guessing on all the raised sentences. That we find exactly one exceptional child is therefore consistent with the mathematical expectations of random guessing in the face of an immature grammar.[80]

In fact, it should strike the reader as rather amazing that the three-year-old who does well on raising *also* does well on passives, and likewise, that the nine-year-old who

[80] One might wonder about the probability of scoring above chance on both raising and passives when neither is actually comprehended. Of the 55 children tested, 26 were found to be above chance on both structures. How many of these cases are expected given random guessing? Zero. [0.073 * 0.0384 = 0.154].

fails on raising *also* fails on passives.[81] While maturation theories predict that there should be few such children (namely, very young children succeeding on these structures, and much older children failing on them), it is quite telling that these apparent age "exceptions" are not exceptions to the correspondence predictions of the maturation accounts. The maturation theories receive tremendous support from the (near) perfect correspondence between the within-subject development and acquisition of raising and verbal passives.

2.4.4 *Conclusion*

Studies 1 and 2 demonstrate delayed acquisition of raised sentences, at least when such sentences include an experiencer. While the timing of acquisition for the raised sentences parallels that found for verbal passive acquisition from prior studies (namely, grammatical development around seven years of age), it was unclear whether delays in acquisition for raising and passives are linked by a common syntactic factor in child grammar. Study 3 helps determine whether such a link exists.

First, similar developmental curves are found for raising and passives, whether measured as a function of accuracy scores across age groups, or as a function of percentage of children within an age group who score above chance. Both curves demonstrate a lack of comprehension (chance performance in the case of accuracy scores, zero or relatively few children in terms of above-chance performance), with no real

[81] It is even more impressive that poor raising and passive scores correlate in the nine-year-old child, since it is known from other experiments in which this child participated that she has for her age above-average IQ as measured by the Kaufman Brief Intelligence Test (KBIT), above-average vocabulary as measured by the Peabody Picture Vocabulary Test (PPVT), and above-average grammatical competency (at least for non-raising and non-passive structures) as measured by the Test for Reception of Grammar (TROG). In light of these facts, poor performance on raising and passives demonstrates that the acquisition of these structures is independent of other aspects of both general cognition and linguistic development, while nonetheless being dependent on one another.

development until about age seven, when comprehension spikes (again, noted in both accuracy and above-chance curves).

Second, individual children's scores on raising and passives are found to correlate significantly and strongly. Above-chance scores are noted as being a better indicator for considering maturation predictions of acquisition, since these theories predict that an individual child will or will not have acquired the syntax that underlies comprehension of both structures (as measured by above-chance comprehension), not that a child's particular comprehension accuracy scores should mirror each other across structures. As such, when above-chance performance across both structures is analyzed on a within-child basis, one finds a near perfect correspondence between comprehension of both structures. The few apparent exceptions are noted to either be very weak exceptions, whose data would not constitute an exception if a single item had been answered differently, or not exceptions at all, once mathematical predictions of chance behavior given a binomial distribution for random guessing is taken into consideration.

These data thus suggest that maturation accounts such as UPR and UFH are on the right track. Namely, StS raising and verbal passive development appear to be strongly linked, both across individual children, and more importantly, within individual children. Such links in development strongly hint at the possibility that a single grammatical deficit underlies the noted delays in acquisition for both structures. Whether the noted delays are due to the reasons put forth by UPR or UFH, or due to causes yet recognized, cannot be determined given the current empirical results. Rather, more data is needed.

Since UPR and UFH differ with respect to their predictions concerning children's comprehension of StS raised sentences without an experiencer-phrase, it would seem that studies investigation raising comprehension without an experiencer would be most welcomed. In addition to being able to help differentiate between two acquisition theories, such data would also help determine whether StS raising is actually delayed, or whether the results on raising obtained in the prior studies were simply tied to the inclusion of the experiencer. If raising is likewise delayed when the experiencer is omitted, one would like to know if raising and passive comprehension continue to mirror each other in development (again, across and within individual children). These issues, and more, will be taken up in Chapter 3.

2.5 *Summary and Subsequent Directions of Inquiry*

We conclude that premature children have (for the most part) no problem comprehending unraised structures, but demonstrate a very significant delay on StS raised structures. Good performance on *think* trials and unraised *seem* rules out "cognitive complexity" explanations for difficulties with raising. The fact that children struggle with StS raising even when the experiencer is fronted suggests that processing accounts working only over surface representations cannot account for children's comprehension difficulties with raised sentences. Near-perfect correlations between comprehension of raising (over an experiencer) and verbal passives strongly favor theoretical explanations that link the acquisition of both structures in development.

2.5.1 *Study 1*

Much has been learned through this inaugural study investigating children's comprehension of unraised and raised sentences. Children are found to comprehend unraised sentences with the raising verb *seem* and an experiencer-phrase. This finding rules out EARH as a proper account of children's difficulties with verbal passives. Children, at least until around seven years of age, fail to comprehend raised sentences with *seem*, when raising takes place over an experiencer-phrase. While poor comprehension of sentences involving StS raising rules out CAH (as does recent data arguing that children know raising to object), the experimental data gathered so far are compatible with both UPR and UFH. The fact that the apparent age of acquisition for raising appears to mirror that found in previous acquisition studies of verbal passives lends further support to UPR and UFH, which are both intended to account for delays in passive acquisition.

Children's particular pattern of comprehension for raised sentences with different foil types, namely chance performance on ER and MR foils, and below-chance performance on DR foils, demonstrates that children employ an interpretive strategy when parsing these raised sentences. Children are not randomly guessing. The *imagine*-analysis, in which children treat *seem* in StS raised sentences with a medial experiencer (predicted to be ungrammatical) as a sort of raising to object verb (predicted to be grammatical), was posited to account for such behavior. The reasons behind why children employ this interpretation, whether it is a compensatory heuristic, a strategy specific to the particular experimental paradigm, or a misanalysis of the lexical item *seem* remain unclear. What is evident, however, is that many children do not possess adult interpretation for the StS raised sentences tested. Whether children's comprehension failures are due to the grammatical problems outlined by either UPR or UFH, or to some other grammatical concern, or even to processing issues, requires further experimentation.

2.5.2 *Study 2*

Here the basic experimental ideas of Study 1 are re-tested, but with slightly different participants and test sentences. The youngest age group and oldest two age groups from Study 1 are not included in Study 2, since the youngest children in Study 1 had such great difficulties with the control conditions for investigating StS raising (namely, the *think* and unraised items), while the oldest children in Study 1 had so few difficulties with any conditions that they proved to represent a solid ceiling for age of raising acquisition, and would thus likely offer little in follow up experiments. This led to only four- to seven-year-olds being included in Study 2. Furthermore, only children who performed at above-chance levels for active transitive, *think*, and unraised sentences were included, so that results on raised sentences could be more faithfully interpreted.

In order to explore the non-grammatical alternative possibility that processing costs associated with integrating elements across an intervening DP explain the data from Study 1, this study has the experiencer-phrase fronted to sentence-initial position for both unraised and raised sentences. With a non-intervening experiencer, at least in the surface representation, children continue to have difficulties with raised sentences. While accuracy on the raised condition is noticeably higher in Study 2 than Study 1, a careful probing analysis by foil type demonstrates that when the relevant foil, in this case the MR foil, is examined, children's comprehension failures with raised sentences are nearly identical between the two studies. In fact, Study 2 mirrors Study 1 in finding acquisition of StS raising to really only occur in the seven-year-olds.

Children's pattern of responses suggests that they are not employing an *imagine*-analysis, which would have suggested chance performance across all foil types. Rather, children appear to be employing either a sort of copular analysis, once they have ignored the experiencer-phrase, or an *expect*-analysis (again, ignoring the fronted experiencer) where the agent is valued over the experiencer. Given that children fail to comprehend the raised sentences even when the experiencer is fronted, a processing account appears untenable as an explanation for children's selective difficulty with StS raised sentences in these studies.

2.5.3 *Study 3*

Certain grammatical theories put forth to account for children's delayed acquisition of verbal passives also expect delayed acquisition for StS raising. Both UPR and UFH predict children to acquire verbal passives and raising over an experiencer at the same time, since on both theories, the relevant grammatical phenomenon accounting for delay is shared between the two structures. While Studies 1 and 2 both find StS raising to be acquired around age seven, the same age when passives are known to be acquired from previous research, it is only with Study 3 that one obtains within-subject data to examine the similarity between developmental curves for these structures, and to ask questions about within-subject correlations between raising and passives in acquisition.

The passive results mimic those found by previous researchers who have crossed voice and verb type, namely, delayed acquisition for subject experiencer passives until around seven years of age. Plotting accuracy scores across age groups for raising and passives reveals strikingly similar patterns of development, with basically chance performance for both structures until the seven-year-old age group, at which point accuracy jumps noticeably and to above-chance level. A similar finding is obtained when individual subject above-chance scores are used in lieu of accuracy. Examination of whether individual children's scores on raising and passives match clearly demonstrates a strong and statistically significant correlation between when the two structures are mastered. While UPR and UFH both correctly predict that verbal passives and subject-to-subject raising structures (with an experiencer) should be delayed, the correlation results from Study 3 serve as further support for both accounts.[82]

2.5.4 *Challenges and New Directions*

While the raising data from Studies 1-3 serve to rule out certain grammatical accounts that had previously been posited in an attempt to explain children's difficulties with verbal passives (e.g. EARH, CAH), and support others (e.g. UPR, UFH), while suggesting that certain processing accounts cannot be right, these data do not definitively suggest why children are delayed in raising and passive acquisition. That is, given these

[82] Hyams and Snyder (2005) write in their Footnote 6: "Note that children need not master passive and [raising over an experiencer] concurrently. This is because learning, as well as maturational 'antifreeze', is required for each of them." As has been acknowledged throughout this chapter, it is clear that the role of learning for successful comprehension of both passives and raising is quite large. That said, the inclusion of particular control conditions within the various experiments (e.g. actional passives for studying verbal passives, the *think* and unraised sentences for studying raising) go a long way in establishing whether or not the requisite learning has already occurred, such that *if* the grammar has sufficiently matured, *then* children should likely succeed in comprehension. In any case, the near-perfect correlations noted between subject-to-subject raising and verbal passives in Study 3 does suggest that Hyams and Snyder might want to reconsider their tepid support for what otherwise seems a clear (and borne out) prediction from their theory.

results, it is impossible to decide between UPR and UFH. Nor do these grammatical accounts themselves have anything to say about the interpretive strategies that children appear to be availing themselves of in Studies 1 and 2.

Further studies are therefore needed that will allow us to gain insight into the particular strategies children employ when attempting to parse StS raised sentences, as well as experimental work to help determine whether UPR or UFH better accounts for children's difficulties with StS raising and verbal passives. In order to accomplish these goals, new studies that manipulate the presence and absence of the experiencer with raising are needed. Within-subject data for raising both with and without an experiencer would help to decide between the predictions of UPR and UFH. Comparing these raising results with passive comprehension, again on a within-subject basis, would further help test the relationship between the grammar underlying StS raising and verbal passive comprehension. Finally, by eliminating the experiencer entirely, we might obtain more insight into the strategies children apply when interpreting raising sentences, since it is possible that different strategies apply when the extra DP is present or absent.

Chapter 3: Comprehension of Raising (no Experiencer)[1]

3.1 More on Comprehension of Raising

In Chapter 2, three acquisition studies (Studies 1-3) were presented that offer compelling evidence that young children, for the most part, comprehend unraised expletive-*it seem* sentences with an experiencer-phrase, but that they do not comprehend the semantically-equivalent subject-to-subject raised sentences (at least not until around the age of seven). Good comprehension of unraised sentences with an experiencer-phrase was taken to demonstrate both that children have learned that raising verbs exist, even if children cannot parse them in an adult manner when StS raising is required, and that experiencer-phrases themselves are not problematic, at least when not appearing in a raised sentence. A processing account working only over surface structures was demonstrated to be unlikely given children also failed to comprehend raised sentences once the experiencer-phrase was fronted.

As far as the specific pattern of results on the raised sentences, various strategies were discussed. To account for StS raising results with both a medial and fronted experiencer, two possibilities were detailed. First, children could apply the raising-to-object *imagine*-analysis when the experiencer appeared medially, and the copula-analysis when it appeared in fronted position. The single *expect*-analysis attempts to explain the

[1] The research presented in this chapter draws in good part from studies first detailed in Hirsch, Orfitelli, and Wexler (2007) and Hirsch, Orfitelli, and Wexler (2008). Study 4 is presented in brief in the former, while Study 5 (minus the control results) is briefly discussed in the latter. Further analyses and significantly expanded discussion is included for both. Study 6 appears in print here for the first time. Once again, all experiments were conducted after careful discussion and planning with Ken Wexler, who also contributed substantially to the interpretation and conclusions. Robyn Orfitelli also helped significantly with experimental design and testing issues.

same data, with the parse switching between a RtO approach with a medial experiencer and a subject-control interpretation without one.

Finally, a strong within-child correspondence was noted between knowledge of StS raising and verbal passives of subject experiencer verbs. Both structures appear to be acquired around the same age (about seven years old), both structures have very similar developmental curves (flat non-comprehension until a spike around age seven), and lastly, statistically significant correlations obtain between scores on both structures.

These results help rule out certain grammatical accounts (EARH, CAH) that had originally been posited to explain children's delayed acquisition of verbal passives. The data, however, are compatible with at least two grammatical accounts of verbal passive delay (UPR, UFH). Both UPR and UFH make similar predictions concerning the ungrammaticality of StS raised sentences with an experiencer-phrase for young children. In order to determine the validity of these two accounts, acquisition studies are needed that would test the comprehension of StS raised sentences without an experiencer-phrase. Before positing any such new studies, let us first consider some relevant data that was unknown when the studies of Chapter 2 were originally conducted.

3.1.1 *Recent Data on the Acquisition of Raising*

Recent experimental work undertaken contemporaneously to that in Chapter 2 by Becker (2005, 2006) bears strongly on the claim that young children do not comprehend StS raising. Becker's central interest in these two papers is in understanding how children come to classify an unknown verb in a DP Verb $TP_{Nonfinite}$ string like (1) as either a subject control verb or a StS raising verb.

(1) Mary *UNKNOW-VERB* to be dancing.

Clearly lexical learning takes place such that a child eventually comes to successful verb classification fitting the frame in (1). Becker concerns herself in part with what types of clues are relevant to making correct inferences about classifying verbs as subject control or StS raising verbs: subject animacy, predicate eventivity, expletive subjects, monoclausal frames, etc. While such lexical issues were not at the heart of the work undertaken in Chapter 2, they are both very important from the general standpoint of language acquisition, and clearly relevant to any discussion about the specific acquisition of StS raising. Crucially, however, Becker also purports to demonstrate that young children not only comprehend StS raising (2), but that they actually analyze (what should be ungrammatical) subject control structures (3) as (grammatical) raising structures.

(2) The hay seems to be on the ground.

(3) #The flower wants to be pink.[2]

These claims come on the basis of two experiments, both of which are presented in Becker (2005) and Becker (2006), just with more child subjects included in the second experiment in the latter paper. In the first experiment, 43 children aged three to five years old are tested on a modified grammaticality judgment task (McDaniel and Cairns, 1990).

---

[2] The pound symbol (#) is used here to indicate ungrammatical *test* sentences, for which an investigator seeks grammaticality judgments. The asterisk symbol (*) is reserved for both sentences adults find ungrammatical and for which grammatical acquisition theories predict ungrammaticality for children.

These children had to listen to a puppet (voiced by the experimenter) comment about a picture they were being shown, and then say whether the puppet's response was "OK" (grammatical) or "silly" (ungrammatical).[3]

The particular sentences uttered by the puppet all consisted of an inanimate subject, a matrix verb, and an infinitival embedded clause. These sentences fell into four conditions, with the experiment consisting of a two-factor design crossing type of matrix verb (subject control vs. StS raising) and "compatibility" between embedded predicate and matrix (inanimate) subject (compatible or incompatible). Two raising verbs (*seem* and *appear*) and two control verbs (*want* and *try*) were used. The embedded clause either contained a predicate requiring an intentional/sentient subject, in which case it would be incompatible with the inanimate matrix subject, or a predicate that did not require an intentional/sentient subject, such that it would be compatible with the inanimate matrix subject. Example test sentences are shown in (4)-(7). In addition to being asked to judge the acceptability of the puppet's response, children were also asked to justify their own responses. The experiment consisted of eight items (two items per condition).

(4)	#The flower wants to be pink.	[Control-Compatible]
(5)	#The flower wants to walk away.[4]	[Control-Incompatible]

[3] The obvious concern with such an experimental design is that it potentially confounds semantic plausibility with sentential grammaticality. I think it is fair to say that young children find it much easier to determine the plausibility of a sentence's meaning than its grammaticality; if children had reliable insights into the grammaticality of their sentences, we would not need to make use of such contrived tasks, and the research in the field of language acquisition would be substantially easier. Alas.

[4] The actual sentence tested in Becker's experiment was *The flower wants to fly away.* Presumably the "incompatibility" of the embedded clause is predicated on the fact that flowers cannot intentionally/agentively fly away. Unfortunately, there exists the non-intentional reading of this sentence whereby a flower might simply float ("fly") away, say due to a stiff breeze, in which case it *is* "compatible" (in the Becker sense) with the inanimate subject. The sentence was changed in (5) above to more clearly demonstrate the intention of Becker's compatibility factor. What effect the ambiguity of the actual test

(6) The hay seems to be on the ground. [Raising-Compatible]

(7) #The hay seems to be excited. [Raising-Incompatible]

Of these four sentence types, adults judge only (6) to be grammatical, while all other conditions are judged ungrammatical. If children have acquired the grammar of subject control and StS raising, and have correctly classified *want* and *try* as control verbs and *seem* and *appear* as raising verbs, they should respond in an adult manner, and should only accept the raising sentences with compatible embedded predicates.[5] If children analyze all test sentences as involving (subject) control, then they should respond that all sentences are silly, on the assumption that children (like adults) require intentional/sentient subjects for control verbs. If, instead, children analyze all test sentences as involving StS raising, then the children should only reject sentences with semantically incompatible embedded predicates.

Becker reports that the five-year-olds she tested respond in an adult manner. The three- and four-year-olds demonstrate a different pattern of results: they respond (basically) correctly to all the raising items (e.g. (6) and (7)), respond correctly for control items with an incompatible predicate (e.g. (5)), but are around chance (36% correct for the three-year-olds, 53% for the four-year-olds) for control items with a

---

sentence had in Becker's study is unknown, though it is likely minor given the mostly adult judgments to items from this condition.

[5] Becker, in her discussions, does not appear to differentiate these two possible sources for children's errors (i.e. lack of grammatical knowledge of control or raising from failure to correctly categorize a verb as a control or raising verb). While these two issues are conflated (e.g. one presumably could not correctly classify a control verb as such without some knowledge that control itself is grammatical), it is nonetheless imperative that any conclusions concerning children's knowledge of grammatical structures not be confounded with children's knowledge of verbs classifications. While these issues are certainly related, they are not the same.

compatible predicate (e.g. (4)).[6] For the moment, let us accept these results as valid (a critical examination is offered in Section 3.3.1), and ask why children might have behaved as such. Becker takes these results to indicate that even young children have acquired StS raising, hence their correct grammaticality judgments on both raising conditions, but have failed to acquire subject control, since they failed on the Control-Compatible condition. Indeed, Becker claims that the three- and four-year-olds are actually misanalyzing the control verbs as raising verbs (no reason is actually given for *why* children would make such errors, or even if subject control itself is ungrammatical or if children have simply failed to correctly categorize control verbs as such). Becker (2005) writes: "This is what I would like to argue children are doing: they initially permit control verbs to have a nonthematic, or raising-like interpretation." Treating the control verbs as raising verbs (the latter which place no intentionality/sentience/agentivity requirements on the matrix subject) means that if the embedded predicate is incompatible with the matrix subject, children would (correctly) reject such sentences, but if the embedded predicate is compatible with the matrix subject, children would (incorrectly) accept the sentences.[7]

Becker does, though, acknowledge the possibility that the children in the youngest two age groups might simply have ignored the matrix verb altogether, parsing only the matrix subject and embedded clause (e.g. for the Control-Incompatible condition getting something like *the flower ... pink*; note, this is somewhat akin to the copula-analysis

[6] Data for the three-year-olds is hardly perfect for the other three conditions, though each was answered at an above-chance level (at a $p<0.05$, but not a $p<0.01$ confidence level): 70% correct Control-Incompatible, 77% correct Raising-Compatible, and 73% correct Raising-Incompatible.
[7] Of course, such a Control-as-Raising account actually predicts *below-chance* (not the noted *chance*) performance on the Control-Compatible condition, but for now, this will be conveniently (for Becker) ignored.

presented in Section 2.3.7). This, Becker claims, would lead to children judging the acceptability of the puppet's responses based solely on the compatibility of an inanimate subject and the intentionality/sentience/agentivity requirements of the embedded clause. Such a strategy would basically result in the noted pattern of results: fine on all conditions except Control-Compatible, since only for this condition is the compatibility of the subject and embedded predicate not sufficient enough information from which to correctly respond.[8]

In order to confront this very possibility, Becker conducts a second experiment. This second experiment is the same in both Becker (2005) and Becker (2006), the only difference being the number of included children (2005: 21 children, 2006: 52 children). The basic findings are essentially identical, but we will concentrate on the specific data from Becker (2006) simply because it includes more data points (more than twice as many children). For the second experiment, only three- and four-year-olds were tested. These children were administered a Truth-Value Judgment task (Crain and Nakayama, 1987). They were told a series of stories (with no pictures employed), and after each story, a puppet commented (truthfully or untruthfully) on it. The children were tasked with having to judge the puppet's response as being either true or false. The sentences the puppet uttered all consisted of an animate subject, followed by either a StS raising verb or a subject control verb, and an infinitival clause. Four raising verbs (*seem, tend, used, happen*) and four control verbs (*want, hope, try, decide*) were tested. The stories were designed, quite cleverly, so that to respond correctly, the child had to parse the matrix (raising or control) verb. Successful performance on this task could thus be used to

[8] The results are obtained if all matrix verbs (both control and raising) are ignored, or if only the control verbs are ignored. Becker's second follow up experiment is ideally suited to determine which, if any, matrix verbs children might be ignoring.

address the issue of whether children were simply ignoring the matrix verbs in Becker's first experiment.

As an example, to test the StS raising sentences, one story scenario involved a white dog which stood under a black (ultraviolet, thus really purple) light, and thus appeared ("seemed to be") purple. The child was then asked to judge the truth of the sentence *The dog seemed to be purple.* Becker (2006) writes, "A child parsing only *the dog . . . be purple* should respond "false", since the dog was not in fact purple; but a child parsing *the dog seemed to be purple* should respond "true" since the dog did seem to be purple when standing under the lamp." Similar scenarios were constructed for the subject control condition. For example, in one, a pig wanted to eat a donut, but actually ends up eating a banana. The child is then asked to judge the veracity of the sentence *The pig wanted to eat the donut.* Ignoring the matrix verb and parsing only *the pig ... eat the donut* should according to Becker lead the child to respond that the sentence is false, since the pig in fact did not eat the donut, whereas the child should respond correctly if he does parse the matrix control verb since the pig indeed did want to eat the donut, regardless of the fact he did not get to eat it.

Both age groups perform at statistically significant above-chance level on both conditions.[9] Based on what appears to be successful comprehension of the control and raising sentences in the second experiment (which crucially require parsing the matrix control and raising verbs), Becker claims that children must be paying attention to the raising and control verbs in the first experiment, and concludes (a) that young children *do*

[9] While both three-years-olds and four-year-olds demonstrate above chance comprehension, it certainly is not the case that they are responding without error, and there is a noticeable (and statistically significant) age effect (Becker, 2006):

Three-year-olds: Raising 64.0% correct, Control 65.9% correct
Four-year-olds: Raising 78.3% correct, Control 88.4% correct

in fact comprehend StS raising structures, and that (b) it is control verbs that they cannot handle, treating them instead as raising verbs. Becker (2006) writes: "[t]he experiments with children show that children do not assume that [a sentence like (1)] has a control structure. If anything, they are inclined to assume it has a raising structure" and "the result [of the first experiment]...is explained not by children's failure to parse the main verb, but by [children's] willingness to assign these sentences a raising structure rather than a control structure."[10]

Accepting, for the moment (see Section 3.3.1 for critical discussion), the validity of Becker's claims (namely, that StS raising is grammatical for young children, and that in fact, even subject control sentences are analyzed as involving StS raising), one is left wondering how Becker's results agree with the potentially incongruent data and claims presented in Studies 1-3. If StS raising is ungrammatical for young children, as proposed by UPR, one must account for what appears to be successful raising comprehension in two other experiments (and even the possibility that subject control involves what had been purported to be ungrammatical StS raising). On the other hand, Becker's data might be the evidence needed to support UFH over UPR, since the former predicts StS raising to be grammatical with no experiencer, but ungrammatical when one is present.

Turning for the moment only to Becker's first experiment, let us consider how her results could fit with the claim that StS raising is ungrammatical for most children younger than seven years old (UPR). Her five-year-olds demonstrate generally adult

[10] Of course, successful comprehension of the subject control sentences in the second experiment basically refutes Becker's claim about how children parse the control sentences in the first experiment. That is, if children did treat control verbs as raising verbs as Becker argues, one cannot account for why children comprehend the control sentences in the second experiment, since if they were being treated as raising verbs, children should have also failed on this second task. Becker offers no comment on these strikingly contradictory results. This, and many other issues with Becker's two experiments, will be critically discussed further in 3.3.1.

judgments for both control and raising sentences. Successful comprehension of subject control sentences offers no challenge to either UPR or UFH, on the assumption that syntactic control is not in the purview of the grammatical claims of either of these accounts as relates to StS raising (cf. Hornstein, 1999, Boeckx and Hornstein, 2003; see Section 4.1.2 for discussion). With regard to the adult-like judgments on the two StS raising conditions, these at first appear to be counter-examples to the claim that StS is not licit in young child grammar (UPR), but are exactly the predictions of an account differentiating StS raising with (ungrammatical) and without (grammatical) an experiencer (UFH).

Before taking these data as evidence that children have acquired StS raising generally (in which case the data in Chapter 2 are unexplained), or even just with no experiencer, one must consider the alternative possibility that such responses might be the product of a compensatory strategy working on otherwise ungrammatical strings (consistent with UPR). The RtO *imagine*-analysis as an explanation here is doubtful given both the inanimacy of the matrix subjects in the test sentences (where RtO verbs require animate subjects), and the lack of a sentence medial DP.[11] Likewise, the subject control *expect*-analysis, which also requires animate subjects, is also unlikely for these raised sentences given that the matrix subjects are inanimate.[12] If children applied a copula-analysis to the raised sentences, however, adult judgments are predicted. The

---

[11] The *imagine*-analysis was posited to account for the raising findings in Study 1, where children appear to interpret StS raised *seem* sentences with a medial experiencer as meaning the same as the also tested *think* sentences. Of course, all matrix subjects in raised sentences in Study 1 were animate. Since RtO verbs require an animate (and sentient) subject, it is extremely unlikely that the *imagine*-analysis could be extended to raising sentences with inanimate subjects (even those with a medial experiencer-phrase).

[12] Given that children appear to have difficulties with the Control-Compatible condition, it is unclear if the *expect*-analysis, which itself requires subject control, is even a possible compensatory strategy for raising sentences with no experiencer. Whether children have by this time acquired subject control is explored in detail in Section 3.3.1.

possibility of a copula-like interpretation was indeed precisely the very reason for Becker's second experiment. Turning to the two younger age groups, the raising results are again captured by the copula-analysis. The subject control data is a bit mysterious, but need not speak to the validity of UPR or UFH, neither of which makes any obvious predictions that subject control should be problematic (again, see Sections 3.3.1 and 4.1.2 for relevant discussions). While Becker takes chance performance on Control-Compatible sentences as evidence for children analyzing them as involving StS raising, this is certainly not a clear-cut interpretation of those results. A detailed critique of this claim is offered in Section 3.3.1, but it suffices to say, if control verbs are analyzed as raising verbs, one expects below-chance as opposed to chance performance, which was not found. Thus, by itself, Becker's first experiment is not incompatible with StS raising being ungrammatical (UPR), once an already-introduced compensatory strategy (i.e. the copula-analysis) is considered.

As far as Becker's second experiment, however, the data offered there, if valid, do present a much more serious argument against a general claim that all StS raising is ungrammatical, and would appear to favor UFH strongly over UPR. Children generally performed quite well on the subject control structures, which does not help differentiate claims about StS raising (though of course, such data quite severely undermine Becker's claim that control verbs are analyzed as raising verbs). If Becker is right, and her test scenarios do require children to actually parse the matrix raising verbs in order to deduce the correct truth judgments (thus ruling out the copula-analysis), then only the *expect*-analysis remains as a possible (as yet previously detailed) strategy. The application of a subject control interpretation to the sentences tested in Becker's second experiment, however, is very unlikely for two reasons.[13] First, unlike in Studies 1, 2, and 3 (and as will be the case in Study 4), the thoughts of the central character are not at issue in Becker's second experiment. All that matters in this experiment is differentiating appearance from reality. Second, even if children attempted to apply such a subject control parse, taking *The dog seemed to be purple* and deriving the meaning *The dog thought he was purple* it is absolutely unclear how such an interpretation could be used to respond (correctly) since the story surrounding this scenario never mentions the dog's belief state as to whether or not he believes himself to be white or purple while standing under the UV light. A crucial question then is whether the data gathered in Becker's two experiments are indeed valid, especially as relates to children successfully responding to StS raising sentences with no experiencer-phrase, a question to which we turn in great detail in Study 5 (Section 3.3.1).

Once again, accepting for now that indeed all of Becker's data *are* valid, a cursory comparison of Becker's two experiments and the two experiments that constitute Studies 1-2 finds certain similarities and differences. Both studies include (minimally) the same StS raising verb *seem* and test children of overlapping ages. The studies do, however, differ in numerous ways, including task methodology, choice of experimental control items, number of test sentences, and potentially crucially, in the inclusion of an experiencer *to*-phrase with StS raising sentences. While Study 2 demonstrates that children have difficulties comprehending StS raising sentences even if the experiencer is fronted, it could be the case that the mere presence of the experiencer is enough to induce

[13] The subject control *expect*-analysis at least now seems (again) viable given children comprehend subject control sentences in Becker's second experiment, and Becker's claim that subject control verbs are interpreted as StS raising verbs appears (again) rather dubious. For evidence that children of the relevant ages have indeed already acquired subject control, see Section 3.3.1.

comprehension difficulties (with raising, not generally, as children have no problem comprehending unraised sentences with either medial or fronted experiencers). Becker's data hint at the possibility, however, that if no experiencer-phrase is present, children successfully interpret StS raising sentences. As discussed in Section 2.2.4, indeed such a pattern of findings would indeed be exactly predicted by UFH, if as Collins (2005a) claims, only raising over an experiencer requires Smuggling (and thus violates UFH's posited ban on exceptions to Freezing). Perhaps then, young children have no problem with StS raising *per se*, but (leaving the details of why aside for the moment) cannot interpret StS raising with an experiencer-phrase, regardless of where it appears in the surface structure. As such, one could account for young children's failure to comprehend the raising sentences in Studies 1-2, which all contained an experiencer-phrase, and their (apparent) successful comprehension of the raising sentences without an experiencer-phrase in both of Becker's experiments.

A further datum that bears on the possibility of the experiencer-phrase being the locus of the grammatical aspect underlying children's failure to comprehend the StS raising sentences in Studies 1-2, comes from the original raising study by Froud, Wexler, and Tsakali (unpublished). These researchers *did* include StS raising sentences without an experiencer (e.g. *Bart seems to be wearing a hat*). Children were found to perform extremely well on such sentences, with both DR and ER foils (accuracy >80% with either).[14] This result is also compatible with the idea that only StS raising over an experiencer is delayed in acquisition, though other plausible alternatives also exist to account for it. For example, this is precisely to be expected if children merely ignored the

[14] Raised sentences with no experiencer were not tested with MR foils since then both the "correct" picture and the foil would be equally valid, as each would show the subject engaged in the action denoted by the embedded predicate.

verb *seem* (copula-analysis) since children can respond correctly just by choosing the picture in which the subject is engaged in the action denoted by the embedded predicate. Also of relevance, none of the pictures used were felicitous with the *expect*-analysis (no reflexive pictures were included). Nonetheless, it does constitute a third example of above-chance comprehension of StS raising with no experiencer.

Clearly then, it is becoming more and more evident that language acquisition researchers might want to separately address the issue of StS *raising over an experiencer* (ROE) from cases of StS *raising with no experiencer* (RNE). This is particularly striking given the (crosslinguistic) grammatical issues noted in the following section.

3.1.2 *Raising over an Experiencer*

We now have to consider the possibility, on the basis of Becker's data, that while ROE appears to be delayed in development, RNE might be early. Given certain assumptions concerning the syntactic operations involved in ROE, some of which are further detailed below in this section, UFH, but not UPR, predicts early grammaticality for RNE and early ungrammaticality for ROE. Interestingly, and *perhaps* of relevance, a similar pattern of grammaticality is also noted in many (adult/natural) languages. In the languages of interest, RNE is grammatical, but ROE is not: Spanish (Torrego, 1996), French (McGinnis, 1998), Icelandic (Thráinsson, 1979), and Greek (Anagnostopoulou, 1997), among many other languages, examples in (8)-(11).

(8) Spanish

*Este taxista$_i$ parece a Maria [t_i estar cansado].

this taxi driver seems to Maria to be tired

'This taxi driver seems to Maria to be tired.'

(9) French

?*Marie$_i$ semble à Jean [t_i avoir bien joué].[15]

Marie seems to Jean to have well played

'Marie seems to Jean to have played well.'

(10) Greek

?*O Jannis$_i$ fenete tis Marias [t_i ekspnos].

The Jannis seems the Marias intelligent

'Jannis seems to Marias intelligent.'

(11) Icelandic

*Hestarnir$_i$ virδast einverjum nanni [t_i vera seinir].

The horses seem some man be slow

'The horses seem to some man to be slow.'

To be clear, in these languages, ROE is not ungrammatical because of any ban on experiencer-phrases or StS raising in general. Here examples from French demonstrate that unraised sentences with StS raising verbs are grammatical with no experiencer (12)

[15] McGinnis (1998), citing her own informants and Rouveret and Vergnaud (1980), notes that for some French speakers, ROE is grammatical, though most find it ungrammatical.

or with an experiencer (13), and where StS raising with no experiencer is also grammatical (14).[16]

(12) Il semble [que Marie danse bien].
It seems that Marie dances well
'It seems that Marie dances well.'

(13) Il semble à Jean [que Marie danse bien].
It seems to Jean that Marie dances well]
'It seems to Jean that Marie dances well.'

(14) Marie$_i$ semble [t_i bien danser].
Marie seems well to dance
'Marie seems to dance well.'

ROE is actually rather rare cross-linguistically. As noted, the ungrammaticality of ROE cannot be attributed to a ban in these languages on raising or experiencer-phrases,

[16] Some languages not allowing raising past a full DP experiencer nonetheless allow raising past a (non-anaphoric) pronoun (e.g. French, Italian, Greek) (i), but others (e.g. Icelandic, Spanish) do not (ii):

(i) Italian (Rizzi, 1986)
?*Gianni$_i$ non gli sembra [t_i fare il suo devere]
Gianni not him seems to do his duty
'Gianni does not seems to him to do his duty.'

(ii) Icelandic (Kim, 2006)
*Ólfaur$_i$ hafði virst þeim [t_i veral gáfaður].
Olaf has seemed them be intelligent
'Olaf seemed to them to be intelligent.'

These data, along with those demonstrating raising past an A'-trace (e.g. topicalization) discussed below, serve to greatly complicate the picture of when ROE is and is not possible crosslinguistically. A full treatment of the data and attempts to capture them is outside the purview of this dissertation. Discussion will be confined to those aspects most relevant to the issues concerning the acquisition of StS raising.

but appears to have something to do with the experiencer serving as an intervener for StS raising. Syntactic movement is generally taken to obey a strict locality constraint, where the phrase that undergoes movement is the "closest" goal to a c-commanding probe. This constraint is captured in the Minimal Link Condition (MLC; Chomsky, 1995) and Relativized Minimality (RM; Rizzi, 2001); see Section 1.2.1. Assuming MLC or RM holds in human grammars, in order to allow ROE, the experiencer-phrase must not c-command (into) the embedded clause. If the experiencer-phrase does c-command (into) the embedded clause, this would entail that the experiencer is actually a closer target for movement to the matrix subject position and, if it is itself a candidate for movement, should move to [Spec,TP]. Otherwise, it should minimally block movement of the lower embedded subject to T. As such, the experiencer intervention effect would really not be that surprising.

The crucial question then is whether or not the experiencer in fact does c-command (into) the embedded clause. Many linguists have noted that there is plenty of evidence that it does (Chomsky, 1986, 1995, Pesetsky, 1995, Kitahara, 1997, McGinnis, 1998). For example, coreference between the experiencer and an R-expression in the embedded clause induces a condition C violation, which is accounted for by Principle C of the binding theory, on the assumption that the pronoun *her* c-commands *Mary* (15).

(15) John$_i$ seems to her$_j$ [t$_i$ to like Mary$_{*j/k}$] [Principle C violation]

As Collins (2005b) notes, all of the c-command tests of the type found in Barss and Lasnik (1986) yield similar results:

(16) The dog seems to every boy$_i$ to like all of his$_i$ toys. [Quantifier binding]

(17) That dog seems to no boy$_i$ to like any$_i$ of his toys. [NPI licensing]

(18) *Who does John seem to who to like? [Superiority violation]

(cf. Who seems to Mary to like who?)

Given the clear evidence above that the experiencer c-commands (into) the embedded clause, the experiencer is actually a closer target for attraction to matrix T than the embedded clause subject, and given MLC or RM, it should raise, but this is ungrammatical in English, whether the preposition *to* is pied-pipped (19) or not (20):

(19) *To Mary$_i$ seems t_i [John is/to be dancing].

(20) *Mary$_i$ seems to t_i [John is/to be dancing].

Attraction of the experiencer to matrix T, however, is grammatical in some languages that do not allow ROE, such as Icelandic (Thráinsson, 1979), with an example given in (21).

(21) Mér$_i$ viðist t_i [Haraldur hafa gert ρetta vel].

Me.DAT seems H.NOM to have done this well

'Harald seems to me to have done that well.'

Movement of the experiencer to matrix subject position, however, is not found in all languages that do not allow ROE. In Italian and French, languages that both prohibit ROE, movement of the experiencer to subject position is also banned, as in the Italian example (22).[17] That is, in such languages, the experiencer blocks the embedded subject from raising, but the experiencer itself also cannot check the EPP feature on matrix T.

(22) *A Piero$_i$ sembra t_i [Gianni fare il sui dovere.]
to Piero seems Gianni to do the his duty
'Gianni seems to Piero to do his duty.'

The evidence is clear that the experiencer c-commands into the embedded clause in English, and on the Uniformity of Theta Assignment Hypothesis (UTAH; Baker, 1988), experiencers in StS raising structures are generated in the same syntactic position crosslinguistically. As such, it is expected under MLC/RM that either the experiencer itself should raise or that, if it is ineligible for movement, it should at least block movement of the lower embedded subject. Examples of each case exist, as noted above, and it appears that ROE, as in English, is actually quite marked crosslinguistically.

How then to account for languages in which ROE is grammatical (what Boeckx (1999) calls the "experiencer paradox")? That is, how can the embedded subject raise when a closer DP experiencer intervenes? Various approaches to this paradox have been

[17] McGinnis (1998) reports that in Italian, unlike French, the experiencer occurs in preverbal matrix position with a finite embedded clause (i):

(i) A Piero$_i$ sembra t_i [che Gianni faccia il suo dovere].
To Piero seems that Gianni does the his duty
'It seems to Piero that Gianni does his duty.'

taken, falling into roughly four classes: late (adjunct) merger of the experiencer, inherent case on the experiencer, Match-Move vs. Match-Agree-Move, and Smuggling.

The first approach involves late adjunct merger of the experiencer-phrase, so that the experiencer c-commands into the embedded clause only after raising of the embedded subject has taken place (Ferguson, 1994, 1996, Kitahara, 1997, Epstein, Groat, Kawashima, and Kitahara, 1998, Boeckx, 1999). As such, these analyses would capture both why the embedded subject can raise, since when it does there is no intervening experiencer, and the c-command facts, since they would hold *after* the experiencer was late merged. There are two important challenges for such accounts. First, the notion of "after" basically entails the existence of two cycles (or counter-cyclic operations), which goes against much work in the Minimalist Program and other strictly derivational approaches. Second, it turns out it is not enough to simply have the experiencer c-command into the embedded clause after raising has been assumed to take place, since there exists evidence that the experiencer c-commands (into) the embedded subject itself, which entrails the experiencer was present before raising took place. On the assumption that there is no reconstruction with *A*-movement (Chomsky, 1995, Lasnik, 1998) then the binding of *himself* in (23) must occur during the course of the derivation, and crucially, before the embedded subject has raised.

(23) [Pictures of himself]$_i$ seem to John [t_i to be ugly].

A second approach revolves around the idea that in some languages, but not all, the experiencer might have inherent Case, and that this inherent Case renders it ineligible for movement (i.e. inactive for attraction), so that the embedded subject may freely raise past it (Chomsky, 1998, McGinnis 1998). Thus, on these accounts, while the experiencer does c-command into the embedded clause once it is merged, it does not block movement of the embedded subject since it has no (structural) Case-feature to match that of the probing matrix T. Two challenges, however, plague such accounts. First, others have argued on the basis of empirical evidence that inherent Case is licensed much like structural Case, that is, in a checking relation (Lasnik, 1995, Collins and Thráinsson, 1996). If true, then it is unclear why the experiencer would not be a possible goal for attraction to matrix T just like the embedded subject. Second, there is evidence that in fact the experiencer is not completely inert/inactive as required on the inherent Case analyses. Boeckx (1999, 2000) argues that while raising of the embedded subject across the experiencer (i.e. ROE) is clearly possible in English (internal Merge is not blocked by the intervening experiencer), subject-verb agreement is blocked with an intervening experiencer. When no experiencer is present, the raising verb shows agreement with the singular or plural post-verbal associate (24). With an experiencer present, however, the raising verb has default (third person, singular) agreement, as seen with the plural associate (25). Thus, at least for those speakers for whom the judgment in (25) holds, the experiencer must be active, as it disrupts an aspect of Agree between the associate and matrix T.

(24) There seem to be men in the room.

There seems to be a man in the room.

(25) There seems/?*seem to Mary to be men in the room.[18]

There seems to Mary to be a man in the room.

Kim (2005) attempts to capture these data by differentiating those languages in which Agree is a necessary precondition for Move from those where Agree need not hold for Move to occur. According to Chomsky (2001), a set of uninterpretable phi-features of a Probe, which is active because of these features, searches for matching features in a Goal, taken to be the closest c-commanded element with the relevant features. If some uninterpretable phi-feature in a Probe matches one in a Goal, they enter the Agree relation, whereby the uninterpretable phi-feature of the Probe becomes valued and deleted. Furthermore, for Chomsky (2001), phrasal movement (Move) requires Match and Agree. Kim's idea is that some languages allow Move without Agree, where feature identity alone is enough to trigger movement. Such is taken to hold of Icelandic, where with a dative ("quirky") subject, agreement is with the object and not the subject (26).

(26) Jón$_i$ virδast/?*virδist t_i vera [Kim, 2006]

John-DAT seem-PL/*-SG to be

taldir/*taliδ t_i lika hestanir.

believed-NOM-MASC-PL/*NEUT-SG to like horses-NOM-MASC-PL

'John seems to be believed to like horses.'

---

[18] As also noted in Boeckx (1999), the intervention effect in (25) seems to disappear where the experiencer is a pronoun (i), the relevance of which is not explored here:

(i) There seem to her to be two men in the room.

For Kim, ROE is possible in languages where movement is done only under Match (no Agree). Since MLC/RM only apply to Agree and not Match, in the case where Agree is not required, the embedded subject is free to raise. If a language also requires Agree for movement, the presence of an experiencer will result in the experiencer (the closest c-commanded DP) entering into Agree with matrix T, such that the embedded subject cannot move.[19] Note that on such an account of ROE as this, UFH predicts that children should have no difficulties with either RNE or ROE. UPR would of course predict problems for both.

The final type of approach to dealing with ROE considered here is that of Collins (2005a), which involves the Smuggling of the embedded subject past the experiencer (see Section 1.3.2 for details). As such, the experiencer c-commands (into) the embedded subject (when the experiencer is first merged via the APPL phrase), but does not serve as a barrier to movement of the embedded subject (since when matrix T probes, the embedded subject has already been Smuggled past the experiencer). Collins is able to capture Boeckx' agreement blocking data, since in those sentences, no Smuggling takes place, and therefore the experiencer indeed is the closer target for matrix T.

The possible relevance of these crosslinguistic data (and related syntactic theories attempting to account for them) is that they raise the possibility that young children simply assume their native language to be one in which ROE is disallowed. That is, the noted delay for ROE from Studies 1-3 might not be due to any grammatical deficit, but

[19] Of course, Kim is left to explain why the experiencer cannot raise in English or French. That is, something in these languages must prevent the experiencer from entering Match with matrix T. Furthermore, Kim's analysis also requires an explanation for why some languages allow movement under Match alone, while others require Match and Agree. Finally, Kim has to explain Icelandic. Kim takes (26) to demonstrate that movement in Icelandic *does not* require Agree, but the lack of ROE in Icelandic (at least with a full DP experiencer) would mean under this account that Icelandic *does* require Agree for movement; else the embedded subject should be able to raise over the experiencer under Match alone.

rather, due to children misanalyzing their own language as one which bans ROE. Under such an idea, children are not subject to non-adult grammatical "rules" (as suggested by maturational accounts like UPR and UFH), but instead assume their language to forbid ROE, for whatever syntactic reasons preclude such structures in adult speakers of Italian and Icelandic.

Presumably the idea is that ROE is somehow "marked", that children start out with the unmarked value (no ROE) of the parameter (assuming a single parameter captures this issue) governing whether ROE is licit in their language, and "learn" or "reset" this parameter with experience. On the Subset Principle (Manzini and Wexler, 1986), children begin with all parameters set to their most exclusive setting. The idea is that by starting with only exclusive settings, only positive evidence is necessary for children to reset the parameters to their correct language-specific setting. While the grammaticality of ROE in any language almost certainly does not simply boil down to an issue of setting a single parameter (especially given the complexity of the cross-linguistic data), nevertheless, entertaining such a (strawman) theory for the moment, many questions would need to be addressed. First, such an account goes up against Borer and Wexler's (1987) "Triggering Problem". Why should it take seven years for the child to reset the parameter, given that the child has the relevant experience (see Section 4.1.1 for evidence that young children in an English-speaking linguistic environment hear plenty of *seem* raised sentences)? Why do most children suddenly set it at exactly the same age (but not sooner)? In other words, even if this were the correct account, we would still need a maturational theory to explain the slow development. This would be even more surprising since learning language-specific parameter settings is (generally) accomplished very early (Wexler's (1998) *Very-Early Parameter-Setting*; see Section 1.1). Lastly, such a parameter setting account would have to explain the strong correspondence between the acquisition of verbal passives and ROE sentences (Study 3), which would appear rather surprising if ROE is only delayed due to failure to set a parameter (since this should have no effect on verbal passives).

It is worth noting, that even *if* such an account could be spun, the exact type of misanalysis on the child's part is rather specific. In many languages that forbid ROE, raising is nonetheless possible if the experiencer itself also undergoes *A'*-movement. So, while Icelandic and French do not allow raising across a full DP experiencer, the data are dramatically different if raising takes place over an *A'*-trace, as seen in French with a topicalized experiencer (27), and Icelandic for a *wh*-moved experiencer (28).

(27) À Pierre$_j$, Valerie$_i$ semble t_j [t_i avoir bien joué].
To Pierre Valerie seems to have well played
'To Pierre, Valerie seems to have played well.'

(28) Hverjum$_i$ hafa hsestarneir$_j$ vurst t_i [t_j vera seinir]?
Who.DAT have horses.NOM seemed to be slow
'To whom do the horses seem to be slow?'

Of the languages so far discussed above, only Spanish bans StS raising over an *A'*-trace (29). These data are particularly relevant to the validity of a parameter setting explanation for children's difficulties with ROE sentences given the findings from Study

2, where young English-speaking children were found to not comprehend StS raising structures with a fronted experiencer. The data from that study demonstrate that the children have certainly not "mis-set" any parameter(s) such that they believe they are speaking a language like Icelandic, French, or Italian, since while those languages do rule out ROE across a full DP (e.g. test sentences in Study 1), they all allow ROE where the experiencer is topicalized (e.g. test sentences in Study 2).

(29) *A Maria$_j$ este taxista$_i$ t_j parece [t_i estar cansado].
To Maria this taxi driver seems to be tired.
'To Maria, this taxi driver seems to be tired.'

The claim would have to become that young English-speaking children have the grammar (as far as ROE is concerned) of (some language like) Spanish. This is actually the null hypothesis on the Subset Principle, since Spanish appears to be the most restrictive language in terms of when ROE is allowed (i.e. no StS raising with the presence of any intervener). The issue then becomes one of demonstrating that early English-speakers assume their grammar to be Spanish-like, for which we know of no relevant evidence. For example, McGinnis (1998) claims that the explanation for why Spanish does not allow StS raising across a non-reflexive clitic pronoun (see Footnote 16) centers around Spanish having clitic-doubling (where doubling of the experiencer clitic renders ROE ungrammatical due to the presence of a phonologically null full DP). Would a parameter setting account of the ROE data in Studies 1-3 thus have to claim that

English-speaking children assume their language to have clitic-doubling? Certainly no relevant evidence exists to suggest this is so.

While there are reasons to doubt both Becker's claim that RNE is early and a parameter (mis-)setting account for the ROE findings in Chapter 2, what is, ultimately, needed are experimental data to directly address children's comprehension of RNE sentences. Such data are crucial to help differentiate between UPR and UFH, since while both can account for the data in the previous chapter (children's failure to comprehend ROE sentences), they make differential predictions as to the grammaticality of RNE in young children's grammar. It is to the gathering of such experimental data that we now turn.

3.2 Study 4: Raising with and with no Experiencer

3.2.1 *Motivations*

To follow up on the studies from Chapter 2, and in order to address the issues raised in the previous section, further experiments are needed. In this study, there are several goals. First, using a new experimental methodology, an attempt will be made to replicate the primary result of Study 1, namely that young children comprehend unraised sentences, but not their semantically equivalent StS raising sentence with a medial experiencer until around age seven. Second, assuming that raising over an experiencer is delayed, evidence will be sought that indeed children make use of a compensatory analysis when attempting to interpret otherwise ungrammatical ROE sentences. In Study

1, children do not appear to guess randomly when presented an ROE sentence, but rather, to interpret such sentences as the also-tested *think* sentences, which is compatible with a RtO *imagine*- or *expect*-analysis. It remains unclear how general such a strategy might be. Third, children's comprehension of StS raising with no experiencer must be addressed experimentally. This is particularly topical given three issues: Becker's (2005, 2006) experimental data suggesting RNE is acquired early (and her more general claim that StS raising is not delayed, and is in fact how children treat subject control verbs), crosslinguistic evidence that ROE is often ungrammatical even when RNE is allowed, and finally, the differential predictions of UPR and UFH concerning the grammaticality of RNE in premature child grammar. Both ROE and RNE sentences will be tested in order to obtain within-subject data concerning the acquisition of both structures. Finally, if children are found to have difficulties with the RNE sentences, those errors will be analyzed to determine if either the subject control *expect*-analysis or copula-analysis is used, both posited in Study 2 in an attempt to account for children's errors in parsing StS raising sentences with a fronted experiencer-phrase (that is ignored).

3.2.2 *Experimental Design*

To investigate the possibility that the presence of an experiencer-phrase was responsible for children's poor comprehension of StS raising sentences in Studies 1-2, this study will directly compare children's comprehension of raising over an experiencer and raising with no experiencer. The methodology to be used to assess children's comprehension of such structures is a Truth-Value Judgment (TVJ) task (Crain and Nakayama, 1987, Crain and Fodor, 1993). This paradigm involves the child participating in a “game” in which he observes an enacted scenario, then listens to an error-prone puppet comment on it, and lastly responds whether or not the puppet commented truthfully or untruthfully. In this study, scenarios consist of interactions between two popular childhood dolls *Barbie* and *Ken*, which are observed by both the child and a stuffed animal named *Mr. Bear.* The child is told that Mr. Bear will offer some comment about the acted-out happenings at the end of each scenario, but that he is somewhat silly and will oftentimes make mistakes. The child is told that he must serve as Mr. Bear’s teacher for the day, and let him know when he is right or wrong, and why.

In each scenario, one or both of the non-puppet characters (Ken and Barbie) says something incorrect about what is occurring; prior to the experiment starting, they are also explained to be quite silly and prone to error. After Barbie and Ken interact, Mr. Bear comments on some aspect of what occurred, using one of four sentence structures: sentences with *think* and a finite embedded clause (30), unraised *seem* sentences with an experiencer-phrase (31), ROE *seem* sentences (32), and RNE *seem* sentences (33), all of which have both a true and false variant (Table 3.2.1). See Appendices (A5) for details.

	True Items	False Items
(30)	Ken thinks Barbie is wearing a hat.	Barbie thinks she is wearing a hat.
(31)	It seems to Ken that Barbie is wearing a hat.	It seems to Ken that Barbie is wearing a hat.
(32)	Barbie seems to Ken to be wearing a hat.	Ken seems to Barbie to be wearing a hat.
(33)	Barbie seems to be wearing a hat.	Barbie seems to be wearing a hat.

Table 3.2.1: Example test sentences from each of the four conditions in Study 4.

To better understand how the scenarios unfold, consider the following example, which is (almost) applicable for test sentences from every condition (Figure 3.2.1). Barbie is wearing her favorite hat. Yet, she does not realize this, and searches for it everywhere, stating as she looks that she cannot find it, but remarks that at least she can see that Ken is not wearing it. Ken, who is standing a good distance away from Barbie, tells the child, away from Barbie hearing, that he believes that he can see that Barbie is wearing her hat. Following the presentation of the scenario, the child is asked three background questions by the experimenter before Mr. Bear comments on what occurred: "Is Barbie wearing a hat?", "Does Barbie know she is wearing a hat?", and "Does Ken know Barbie is wearing a hat?". These questions serve to establish that the child understands all of the relevant details of the scenario, and help rule out many non-grammatical reasons for any possible response errors. If the child does not answer these background questions correctly, the scenario is simply repeated.

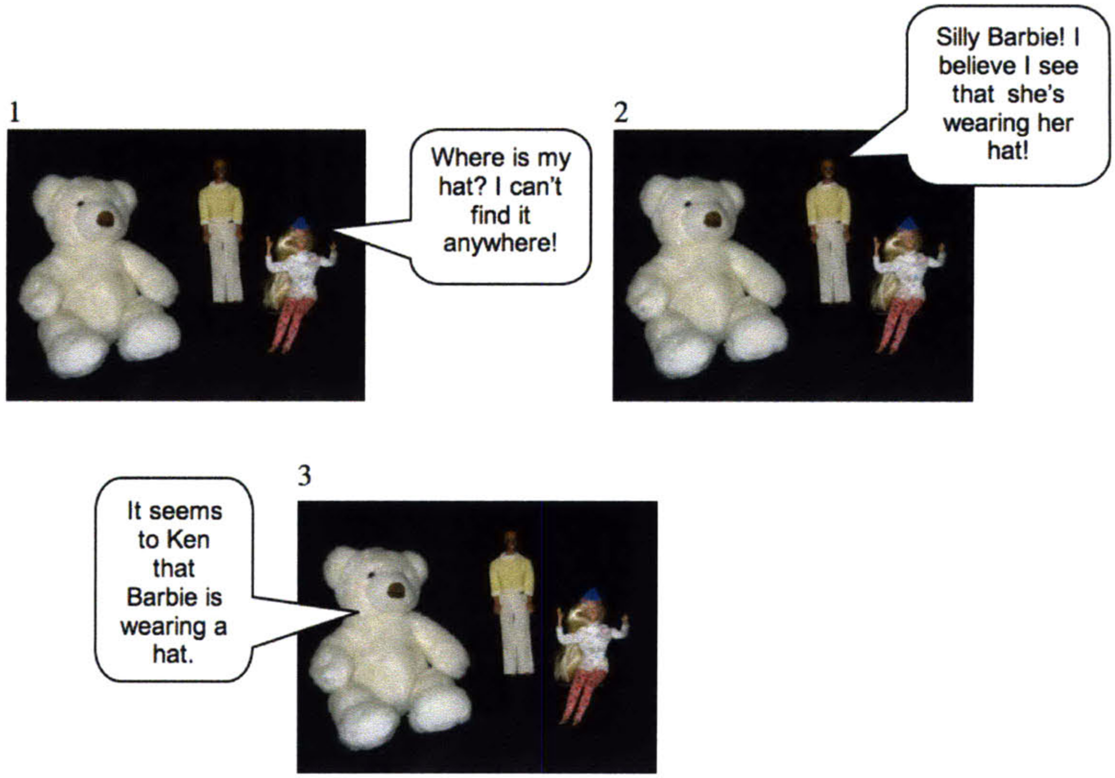


Figure 3.2.1: Example scenario from Study 4 in which (1) Barbie proclaims that she cannot find her hat (though it sits unbeknownst to her on her head), after which (2) Ken comments, away from Barbie and to the child, that he believes Barbie is wearing her hat, and finally, (3) Mr. Bear utters a test sentence (here, a True unraised condition item).

The *think*-condition (30) is meant to serve in part as an attentional control, establishing that the child understands and is paying attention to the task. Responding successfully requires children to follow complex interactions between the two characters, paying attention to their dialogue and object manipulations, plus the commenting puppet. Children also have to be able and willing both to respond to the puppet *and* justify their answers. Since crucial manipulations within the task involve tracking the belief state for each of the three puppets, this TVJ task also requires children to have developed a strong sense of Theory of Mind (ToM; Wellman, Cross and Watson, 2001). To do well on the *think* sentences, likewise for the other three conditions, children must be able to understand verbs and scenarios requiring knowledge about other people's beliefs.

The task has been designed such that in order to respond correctly to the *think*-condition sentences, children cannot simply parse the embedded clause (i.e. ignore the matrix subject and matrix verb *think*). To do so would lead to consistently below-chance performance on the False trials. For example, in the scenario described in Figure 3.2.1, given the False *think* sentence *Barbie thinks she is wearing a hat*, if a child only parses the embedded clause *she is wearing a hat*, he should respond "true" since Barbie is indeed wearing a hat. Barbie, however, clearly states that she does not know where to find her hat, so the correct response to this test item is "false" as Barbie does not think she is wearing a hat. As such, not only is this condition testing a child's ability to understand false beliefs, but it also serves to establish that the child can interpret grammatical structures with embeddings. Furthermore, it helps pick out those children unwilling to state that the puppet is wrong, which is crucial since some children when administered a TVJ task cannot bring themselves to say that the puppet (or experimenter) is ever incorrect (a sort of "yes"-bias). Therefore, only those children who successfully (defined below) respond to the *think* sentences will be included for subsequent analyses.

Turning to the unraised and two raised conditions, following Study 1, the verb *seem* will serve as the only StS raising verb in this study. As noted there, *seem* is not only the most frequently used StS raising verb with an unraised variant in child-directed speech for the English portion of the CHILDES corpus, but it is also frequent in absolute terms (both in its unraised and raised forms), appearing more often than verbs such as *dance*, *crawl*, and *hug*; see Section 4.1.1 for details. Thus, using *seem* minimizes the possibility that any problems children might have with StS raising are due to not having heard the verb or to not having heard it in its raised form. Furthermore, choosing a StS raising verb like *seem* that has an unraised variant helps to ensure that the child knows the verb, so that any failures with StS raising can be differentiated from simple failures of StS verb learning. Lastly, by again only using *seem*, the results from this study can be directly compared to Studies 1-2 where the same StS raising verb was used.

To ensure that any observed poor performance on the two StS raising conditions is not simply due to children not knowing the raising verb *seem*, an *unraised* expletive-*it* condition with *seem* is included (31). Importantly, the false items of this condition establish that the child actually understands the raising verb, and does not merely parse the embedded clause to arrive at the correct answer. Given the scenario described above, in which Barbie does not realize that she is wearing a hat, a child could correctly identify that *It seems to Ken that Barbie is wearing a hat* is true by solely attending to the embedded clause. To keep the child from being able to ignore the matrix verb (*seem*) and experiencer, and still respond correctly, the false items for this condition had slightly altered scenarios. For the scenario in Figure 3.2.1, the false test item would involve the following: Ken stands so far away from Barbie that he cannot see her very well, and thus (mistakenly) concludes that she must be wearing a scarf. Barbie, however, is wearing a hat and not a scarf (just as in the true item), so the embedded clause of the test sentence *It seems to Ken that Barbie is wearing a hat* is true, but because Ken believes that she is wearing a scarf, the test sentence is actually false. A child only attending to the embedded clause would therefore incorrectly answer "true" to these items. Correctly responding to the false unraised condition sentences therefore establishes that children comprehend the verb *seem* with an experiencer-phrase, at least in its unraised form. Furthermore, as one of the major goals of this experiment is to establish whether or not it is the experiencer-phrase that causes children's difficulties with StS raising structures, good performance on this condition also establishes that it is not the mere presence of an experiencer-phrase in a sentence that causes difficulties for children on raising items.

The raised conditions, with (32) and without (33) an experiencer-phrase, constitute the crux of this study. While UPR and UFH both predict children will be delayed in comprehending ROE sentences, they differ in their predictions regarding RNE sentences. UPR predicts children will also be delayed on these latter structures, while UFH predicts no such delay. Note that Becker, having concluded that StS raising is not delayed, predicts children should have no difficulties with either raised condition, nor of course does her account offer an explanation for the findings of Studies 1 and 2.

The scenarios are specifically designed such that a child using either the RtO *imagine*- or *expect*-analysis (both meant to capture the semantics of a *think* sentence) should consistently answer all of the ROE sentences incorrectly.[20] Recall from the scenario described above that while Ken thinks Barbie is wearing a hat, Barbie does not realize that she is, and explicitly states that she can see that Ken is not wearing a hat. Upon hearing the true ROE sentence *Barbie seems to Ken to be wearing a hat* then, a child subject to either of the RtO analyses will incorrectly answer "false" because Barbie

[20] To be clear, the *imagine*-analysis and *expect*-analysis are exactly same when applied to ROE sentences. Basically, they involve RtO syntax with the semantics consistent with the (non-RtO) verb *think*. That is, the matrix subject is taken to be an experiencer (of the "thinker" type), while the embedded clause serves as the content of the thinker's thought. This is quite compatible with the verb *imagine*, but a little less evident with the verb *expect* (which expresses "expectation" over "thought"). Again, using the label *expect* for the analysis is only meant to capture the fact that *expect* has both a RtO analysis (with medial DP) and a subject control analysis (without medial DP). The semantics in both cases is supposed to match that of *think*; unfortunately there is no such exact English verb that would make naming the analysis more clear.

does not think Ken is wearing a hat. Likewise, for the false item *Ken seems to Barbie to be wearing a hat*, a child will again respond incorrectly ("true") if using either RtO analysis because Ken does think Barbie is wearing a hat. Thus these analyses, whereby children treat the ROE sentences as having the meaning of a *think* sentence, lead to consistently below-chance performance on this condition, as both true and false items are answered incorrectly. That Mr. Bear uses a *seem* sentence in these cases, which reflects some doubt on the part of the speaker that is not expressed in a similar *think* sentence, is precisely licit given Mr. Bear's general sense of uncertainty in any of his answers due to his silliness.[21]

In addition to including ROE sentences, this study seeks evidence concerning children's comprehension of RNE sentences. As proposed in Study 2 in an attempt to account for children's comprehension of StS raising sentences with (what appears to be) an ignored experiencer-phrase, leaving basically a RNE sentence (34), children could be applying either a copula-analysis or subject control *expect*-analysis. The copula-analysis has children ignoring the raising verb and substituting in its place the copula (35).[22]

(34) ~~To Mary,~~ John seems to be dancing.

(35) John is dancing.

[21] That the use of *seem* sentences in this task is perfectly acceptable is generally confirmed on the basis of both adult subjects who took part in pilot testing and who expressed no issues with their inclusion, and the very good performance on these sentences by the tested seven year-olds (3.2.5). See Footnote 24 for similar comment concerning the RNE sentences.
[22] The exact syntactic details, as relates to the *to be* material in the embedded clause, are left unaddressed (e.g. in (34) whether only *seem* is ignored leaving *John to be dancing* which becomes *John is dancing*, or whether *seem to be* is simply replaced with the copula to produce directly *John is dancing*).

While this is certainly an intriguing possibility for children who might lack StS raising, it is not the primary strategy that is to be explored in this study. Rather, the scenarios and test sentences are designed to investigate the possibility that children lacking StS raising syntax make use of the *expect*-analysis. This analysis is meant to unify (to a degree) the interpretations children might derive for ROE and RNE sentences. The idea is that a single lexical item is interpreted in place of the StS raising verb, with RtO syntax in the case of ROE and subject control syntax with RNE, where both express "thinker" semantics. Thus, when premature children hear the ROE sentence in (36), which is ungrammatical on the adult (StS) analysis for them, they interpret it as (37). Similarly, when they hear the RNE sentence in (38), they interpret it like (39). Note, though, that while the verb *expect* is being used here (precisely since it is compatible with both structures), the actual meaning intended is one with semantics closer to *think*.

(36) John seems to Mary to be dancing.

(37) John expects Mary to be dancing.

(38) John seems to be dancing.

(39) John expects to be dancing.

I assume it is just a fluke, however, that there is no such verb in English with this exact meaning and these exact syntactic properties. The verb *claim* might be a better example than *expect* (as suggested by C. Schütze, personal communication), but even that is not quite right semantically. That subject control verbs exist with a meaning of *think*, however, is clearly demonstrated by the existence of just such verbs in other languages, as seen in the French (40) and German (41) examples below.[23]

(40) Jean pense porter un chapeau. [French]

Jean thinks PRO to wear a hat.

'Jean$_i$ thinks he$_i$ is wearing a hat.'

(41) Franz denkt, einen Hut zu tragen. [German]

Franz thinks PRO a hat to wear.

'Franz$_i$ thinks he$_i$ is wearing a hat.'

The crucial aspect of the *expect*-analysis when applied to RNE sentences is that it results in reflexive meaning. Specifically, the sentence is one in which the matrix subject thinks he himself is doing the action denoted by the embedded predicate. To see how this might apply to the current study, consider the scenario described previously in which Barbie is wearing a hat but does not know she is. Mr. Bear is perfectly correct to state (the true RNE sentence) *Barbie seems to be wearing a hat.*[24] Yet if a child could not comprehend this sentence due to the involved StS raising syntax, and were to apply a subject control *expect*-analysis (again, with semantics expressing "thinking" not

[23] The French and German examples, however, have no RtO equivalent. In any case, the idea of the *expect*-analysis is that, at least with animate subjects, children analyze *seem* in ROE as a RtO verb with semantics similar to *think* and in RNE as a subject control verb with semantics similar to *think*.

[24] As noted for the case of the ROE sentences, using a *seem* RNE sentence (e.g. *Barbie seems to be wearing a hat*) here, say as opposed to a copula sentence (e.g. *Barbie is wearing a hat*), does express some doubt about the surety of the response by the speaker. Given Mr. Bear's silliness and general uncertainty, which is made clear to the child before the experiment begins, this use of *seem* is quite natural. To support that claim, note that adults who took part in a pilot version of this experiment expressed no reservation about the use of RNE sentences in this context, nor do older (seven year-old) children have any problem providing adult judgments of these sentences (see 3.2.5).

"expecting") to this structure, he would consistently answer incorrectly ("false"), since it is not the case that Barbie thinks that she is wearing a hat (she explicitly states that she has no idea where her hat is). The same works with the false items in this condition, though the scenarios are necessarily slightly different. For example, in one such scenario, Barbie believes she is wearing her favorite belt, but in fact, she forgot to put it on. Ken meanwhile can see that she is not wearing a belt, but he does not share this with Barbie. When Mr. Bear comments (falsely) *Barbie seems to be wearing a belt*, a child using the *expect*-analysis, would arrive at a meaning *Barbie thinks she is wearing a belt* (which is true given the noted scenario), and would thus respond "true" where an adult responds "false". One therefore predicts that a child using the *expect*-analysis for raising structures without an experiencer will demonstrate consistent below-chance performance, interpreting the sentences as involving subject-control, giving them a reflexive interpretation, and thus answering all RNE questions incorrectly.

Now consider a child whose grammar lacks the means of representing a RNE sentence, but who instead of applying the *expect*-analysis uses the copula-analysis when presented a RNE sentence. When asked to judge the truth of *Barbie seems to be wearing a hat* for the scenario described in Figure 3.2.1, he will respond (correctly) "true" since Barbie *is* wearing a hat. For the above (false item) case with Barbie incorrectly believing she is wearing a belt, the child will judge (again, correctly) *Barbie seems to be wearing a belt* as "false" since Barbie is not wearing a belt. That is, given the scenarios and RNE test sentences in this study, there is no way based only on children's truth-value judgments to differentiate adult StS raising comprehension from lack of StS comprehension and application of the copula-analysis.[25] From the discussion of Study 2, it was noted that since no reflexive foils were present in that experiment, the copula-analysis was the more likely explanation for understanding children's responses to the StS raising sentences (once the fronted experiencer was ignored). To the extent that the results from Study 2 do demonstrate the application of the copula-analysis, it certainly seems quite possible that the same will also occur here in Study 4. Indeed, errors on RNE sentences are really only expected in the case that children lack StS syntax *and* precisely adopt the subject control *expect*-analysis. As such, interpreting correct responses to RNE sentences in the face of incorrect judgments on ROE sentences must be undertaken judiciously.

If Becker is correct, and children have no trouble with StS raising syntax, we expect children to succeed on both ROE and RNE conditions in this study. If, however, UPR or UFH is correct, then younger children should (at least) fail to comprehend the ROE sentences. In the case of UPR, the same children are also expected to fail to comprehend RNE sentences (whether comprehension errors will be detected depends entirely on whether children make use of the *expect*-analysis, in which case errors should be noted, or of the copula-analysis, in which case errors will go undetected). According to UFH, while ROE is ungrammatical (requires Smuggling), RNE should be fine (need not or does not require Smuggling).

Every child is presented eight sentences from each of the four conditions, with four each of true and false items, for a total of 32 items. The use of eight items per condition, as opposed to only two as used in Becker's (2005, 2006) first experiment, will

[25] When Study 4 was first conceived and carried out, Study 2 had not been undertaken, and as such, the copula-analysis had not yet been put forth. At the time Study 4 was put together, only the *expect*-analysis as a means of dealing with RNE sentences was being considered.

help facilitate the deduction of particular response patterns (e.g. above-, at, or below-chance comprehension levels). These items are pseudo-randomly balanced such that children are not tested on the same condition more than twice in a row. After asking the various background questions following each scenario, the experimenter (now acting as Mr. Bear) reads each test sentence aloud twice before the child is allowed to respond. The child's response justifications are noted before moving to the next test item. Sessions with each child are transcribed in real time and also recorded onto audiocassettes. Taping of the testing sessions is done in order to preserve the detailed child justifications, in case they are too long or complex to be fully transcribed by the experimenter in the daycare setting; in such cases, the experimenter can simply return to the tapes for any missing responses once back in the laboratory. By having children give justifications for their responses, we can attempt to gain greater insights into any errors and reasons for such. In particular, an examination of the justifications, especially for incorrectly answered items, might help establish a deeper sense of any linguistic strategies being employed.

After introducing the characters, and explaining how the "game" was to be played, but before beginning the testing session, children are required to answer three practice items. These were designed in order to confirm that the children understood the general task demands and were willing to correct Mr. Bear when he made false statements. The first practice item had Mr. Bear correctly stating the sex of either one of the experimenters or one of the characters (e.g. Ken is a boy) to which children had to tell Mr. Bear he was right. The second item had Mr. Bear incorrectly stating the sex of someone, either one of the characters or one of the experimenters (e.g. Barbie is a boy, Jeremy is a girl), requiring the children to let Mr. Bear know he was wrong. The third practice item was actually a False *think* sentence, which was included for two primary reasons. First, it ensured that the children could follow the complexity of a real test scenario, including the false belief of one of the characters. Second, since successful comprehension of the *think* condition is to be used as a prerequisite for inclusion in the study, as discussed below, it only makes sense to attempt to distinguish early in the testing session which children might not succeed before wasting both the child's and the experimenter's time, especially since testing took on average 30-45 minutes per child.

As detailed in 3.2.3, in order to ensure that performance on the raised conditions accurately reflects a child's level of grammatical comprehension, as opposed to orthogonal issues having to do with task overload, understanding false beliefs, or lexical learning, children not scoring above chance (less than 7 out of 8 items correct) on each of the *think* and unraised conditions would not be included in any of the data analyses. Instead, they would be replaced until the requisite number of children per age group was reached.

3.2.3 *Participants*

As mentioned in the previous section, those children unable to score at least 7 of 8 correct on both the *think* and *unraised* conditions were eliminated in order to ensure that those children whose data is included comprehend and can follow the task demands, have acquired Theory of Mind (i.e. understand false belief), have no trouble with experiencer-phrases generally, and know the raising verb (*seem*) to be used in the StS raising conditions. This was done to eliminate all factors that might negatively impact performance on raising items other than comprehension of the syntax of StS raising itself.

As was the case in Study 2, only children aged four years old to seven years old are included in Study 4. This is done to maximize testing resources, since even younger children were noted to often not comprehend *think* and unraised sentences in Study 1, while older children all tended to comprehend ROE sentences. Focusing on this particular age range narrows in on the time period in which ROE appears to be acquired by the great majority of children. Not including three-year-olds, however, leaves out an age group Becker tested in her experiments, and thus three-year-olds will be examined in Studies 5-6.

Given the requirement that children successfully respond to the *think* and unraised conditions, certain children, especially younger ones, were tested and subsequently eliminated before reaching a total of 40 children (10 per age group) for inclusion. While no strict numbers were kept on how many children tried but could not complete the task or had to be eliminated due to poor *think* or unraised scores, a rough estimate would be in accordance with the data from Study 1 as to percentage of children not knowing *think* and unraised sentences. In any case, in relatively short order, 40 children were found whose data is worth considering. These 40 children (12 boys, 28 girls) consisted of 10 in each one-year interval between four and seven years of age (4.35-7.97 years, mean age 6.06).[26] Participant details are shown in Table 3.2.2. All included children were normally developing native-English learners and came from families of varying socioeconomic status.

Age Group	#	Mean Age	Youngest	Oldest	Male	Female

[26] The inclusion of over twice as many female children reflects nothing more than the sex makeup of the daycares available at the time of testing. There was certainly no obvious sex difference in terms of which children successfully comprehended the prerequisite *think* and unraised items.

4	10	4.67	4.35	4.97	2	8
5	10	5.42	5.07	5.90	3	7
6	10	6.51	6.08	6.99	3	7
7	10	7.62	7.29	7.97	4	6
Total	40	6.06	4.35	7.97	12	28

Table 3.2.2: Child participant details for Study 4.

3.2.4 *Results*

Accuracy results for all children across the four conditions are included below in Table 3.2.3.

Age Group	Think	Unraised	ROE	RNE
4	96.3%	97.5%	5.0%	40.0%
5	98.8%	97.5%	8.8%	51.3%
6	100.0%	96.3%	40.0%	78.8%
7	100.0%	100.0%	66.3%	87.5%
Total	98.8%	97.8%	30.0%	64.4%

Table 3.2.3: Accuracy for all conditions across all age groups in Study 4.

In Table 3.2.4, the percentage of children in each age group scoring at a statistically significant above-chance level is shown for each condition.[27]

Age Group	Think	Unraised	ROE	RNE
4	100.0%	100.0%	0.0%	40.0%
5	100.0%	100.0%	0.0%	40.0%
6	100.0%	100.0%	30.0%	70.0%
7	100.0%	100.0%	70.0%	90.0%

Table 3.2.4: Percentage of children in each age group performing at above-chance level for all conditions in Study 4.

[27] Calculating above-chance performance using a 95% one-tailed confidence interval for eight items and random two-choice guessing (0.5) gives: roundup[4 + (1.65 * [sqrt(8 * 0.5 * 0.5)]) = roundup[4+2.33] = roundup[6.33] = 7 items minimum. Details concerning how such calculations are derived can be found in Chapter 2 (Section 2.2.3).

On *think* and unraised trials, children perform extremely well, in accordance with the stipulation that any child scoring less than 7 of 8 correct on either condition be eliminated from the study. These high accuracy scores indicate that these 40 children all understand the task, have Theory of Mind, including knowledge of false belief, understand the meaning of *seem*, and can comprehend experiencer-phrases, at least in unraised sentences. We have therefore helped to ensure that children's performance on StS raising sentences is independent of these factors.

Turning to ROE, four- and five-year-olds clearly do not comprehend these sentences (neither age group as a whole is above 10% accuracy). Even the six-year-olds do not score above chance (50%) as a group. It is only with the seven year-olds that accuracy scores jump well above chance level. Taking all 40 children as a single group, accuracy for ROE sentences is only 30.0%, while accuracy for the semantically equivalent unraised sentences is a very impressive 97.8%.

These accuracy results are mirrored in the percentage of children in each group scoring at an above-chance level. Not a single four- or five-year-old comprehends (scores significantly above chance on) the ROE sentences. Indeed, overall, only 25% (10/40) of the tested children score above chance on ROE. Of those children younger than seven years-old, only 10% (3/30) comprehend ROE, while the good majority of seven-year-olds, fully 70% (7/10), comprehend ROE sentences. Thus, the large increase in accuracy at the age group level noted for the seven-year-olds, as compared to the other age groups, is also reflected in the jump in the number of children scoring above chance noted for the seven-year-olds. This apparent late acquisition of ROE using a TVJ task, around age seven, is exactly what was found in Studies 1-2 using sentence-picture matching tasks, and is also right around the age when verbal passives have been argued to be acquired (Section 1.2.2).

Examination of the response patterns of the 75% of children tested who did not score above chance on the ROE condition reveals that it is not the case that these children are merely randomly guessing, or that they manage to get all the true items correct while missing all the false items, or vice versa. Rather, those children who do not comprehend ROE sentences tend very strongly to consistently respond incorrectly to all ROE items, thus demonstrating below-chance comprehension (Table 3.2.5).

Age Group	Chance Type	ROE
4	BC	9
	C	1
	AC	0
5	BC	8
	C	2
	AC	0
6	BC	5
	C	2
	AC	3
7	BC	3
	C	0
	AC	7

Table 3.2.5: Breakdown of chance performance for ROE condition across age groups in Study 4 (BC=below chance, C=chance, AC=above chance). The vast majority of children who are not AC respond at BC level, not C level.

Of the 20 four- and five-year-old children, as already noted, none perform above chance on this condition. Strikingly, however, 85% (17/20) of these younger children score at below-chance level on the ROE sentences. Of the 30 children who do not score above chance on the ROE condition, fully 83.3% (25/30) score below chance. Thus, not only do children younger than about age seven not comprehend ROE sentences, they appear to consistently interpret them to mean something completely different (see following discussion section for elaboration).

Shifting focus, we find that just as in the ROE condition, children perform poorly on the RNE condition, not demonstrating appreciably better than chance performance as a group until around the age of six. Both the four- and five-year-olds, as groups, have comprehension accuracies hovering right around chance level (40.0% and 51.3%, respectively). Across the 40 children, accuracy on the RNE condition is only 64.4%, and only 60% (24/40) of the children tested score at an above-chance level. Of those children younger than seven years-old, only half (15/30) comprehend the RNE sentences. Thus, it is clear that RNE sentences, at least for a very great many children, are not grammatical (on a StS raising analysis; i.e. in the adult sense).

That said, it is also quite clear that RNE comprehension is not nearly as poor as that for ROE sentences. Just looking at overall accuracy on the two conditions for all children, RNE sentences are answered correctly twice as often as ROE sentences (64.4% vs. 30.0%). While only 25% (10/40) of the children demonstrate above-chance comprehension on the ROE condition, 60% (24/40) comprehend the RNE sentences. These findings are even more striking when considering the youngest half of the children (those aged four to five). Not one of these younger children (0/20) comprehend the ROE sentences, but 40% (8/20) respond correctly to the RNE sentences. While RNE performance is generally quite poor, especially compared with the noted near-perfect comprehension of *think* and unraised sentences, comprehension for StS raised sentences with no experiencer is notably better than comprehension for the same sentences with an experiencer. Determining what this difference means will be taken up in the discussion section below.

For the moment, consider the chance type behavior patterns for RNE sentences, as seen in Table 3.2.6.

Age Group	Chance Type	RNE
4	BC	6
	C	0
	AC	4
5	BC	4
	C	2
	AC	4
6	BC	1
	C	2
	AC	7
7	BC	1
	C	0
	AC	9

Table 3.2.6: Breakdown of chance performance for RNE condition across age groups in Study 4 (BC=below chance, C=chance, AC=above chance). The vast majority of children who are not AC respond at BC level, not C level.

As seen in the above table, 16 of 40 tested children do not score at a statistically significant above-chance level on the RNE condition. Of those children who do not comprehend RNE sentences, 75% (12/16) score at a statistically significant below-chance level (at most one item correct) on this condition. Comprehension of RNE sentences therefore appears to be an all-or-none issue for each child. Almost all children score either at above- or below-chance level. Only 10% (4/40) of the children who participated demonstrate chance performance. If children were guessing randomly on the RNE sentences, where the probability of choosing the correct picture in the two-choice sentence-picture matching task is 0.5, given a binomial distribution, 93.7% would be expected to be at chance (i.e. 2-6 items correct). This general lack of chance performance is seen in all age groups; increasing group accuracy is due to higher numbers of children scoring above chance, not to individual children scoring slightly better in each subsequent age group. For those children who do not comprehend RNE sentences, the vast majority are not simply guessing when presented such sentences. Instead, most of these children are consistently providing an incorrect interpretation to RNE sentences. What this interpretation is awaits discussion in the following section.

Before turning to that discussion, though, consider the data in Table 3.2.7. What this table shows is the number of children who comprehend RNE sentences, but fail to comprehend ROE sentences (first column), and, vice versa, those children who comprehend the ROE sentences, but not the RNE sentences (second column). While there are plenty of children who fail to comprehend both conditions, and even a few (mostly older) children who do comprehend both conditions, when these children only comprehend one sentence type, it is always the RNE sentences that are understood. While there are 14 children who comprehend RNE but not ROE, there is not a single child who understands the ROE sentences but not the RNE ones. This is strong evidence that ROE and RNE sentences are not comprehended independently of one another. Rather, there is a clear acquisition directional (but not bi-directional) dependency in effect, namely, ROE sentences are only comprehended by those children who already comprehend RNE sentences.

Age Group	RNE, ~ROE	ROE, ~RNE
4	4	0
5	4	0
6	4	0
7	2	0
Total	14	0

Table 3.2.7: Number of children who score above chance on RNE condition but not ROE condition (first column) and number of children who score above chance on ROE condition but not RNE condition (second column) across all age groups for Study 4.

3.2.5 *Discussion*

In designing this experiment, we set out to address three general issues. First, an attempt was made to replicate the central findings from Study 1 using a different experimental methodology: that while unraised sentences with an experiencer-phrase are generally quite well understood by young children, their semantically equivalent StS raised forms are not comprehended, at least until around seven years of age. Second, comprehension of RNE sentences was investigated, both within the same study and for the same children as tested on ROE sentences. This was done to address both the claim that StS raising is not delayed in acquisition, and that only StS raising over an experiencer is delayed. Third, scenarios and test sentences were construed to better address the validity of the various interpretive strategies proposed for children's analysis of raised sentences in Studies 1-2.

Turning first to the issue of replicating the main findings from Study 1, the current study does appear to validate those earlier data. Given the requirements for inclusion, children again did very well on *think* and unraised conditions, even with the completely new experimental paradigm (TVJ vs. picture selection). The unraised condition is important for two reasons in particular. By requiring that children demonstrate above-chance performance on the unraised condition, it greatly reduces the likelihood that poor performance on the raising conditions could simply be due to children not comprehending the raising verb itself. Also, successful comprehension of the unraised condition also shows that children are not merely parsing the embedded clause (which would result in incorrect responses to false trials), and furthermore, that the mere inclusion of an experiencer-phrase is not grammatically problematic.

As for the ROE condition, once again, children are greatly delayed in their comprehension of StS raising when an experiencer-phrase is included. Comparing comprehension of such structures across Studies 1, 2, and 4 shows remarkable similarity both in the percentage of children who understand them at each age group, and the fact that these structures appear to be acquired around age seven, regardless of the specifics of the experimental task (Table 3.2.8). Across all three studies, four-year-olds are extremely poor in their comprehension of StS raising with an experiencer-phrase (at most 20% comprehend them), and less than one in three five-year-olds understand these sentences. It is only with the seven-year-old groups that a great majority (minimum two in three) of children in an age group demonstrate above-chance comprehension. These data are as predicted by both UPR and UFH, but unexpected if StS raising is not delayed as claimed by Becker.

Age Group	**Study 1**	**Study 2**	**Study 4**
4	16.7%	20.0%	0.0%
5	12.5%	26.7%	0.0%
6	25.0%	40.0%	30.0%
7	66.7%	80.0%	70.0%

Table 3.2.8 Percentage of children per age group who perform above chance on ROE sentences in Study 1, Study 2, and Study 4.

With respect to RNE sentences, here too children demonstrate poor performance. Comprehension across all the children does not exceed 65% accuracy. The good majority (60%) of four- and five-year-olds do not comprehend the RNE sentences. For such children, clearly RNE, just like ROE is delayed (for evidence that these two structures are not independently acquired, see Table 3.2.7 above). It is StS raising, not raising over an experiencer-phrase that is problematic.

Those children failing to comprehend RNE sentences provide strong evidence against UFH. To the extent that UFH follows Collins' (2005b) account of raising whereby raising with no experiencer does not involve Smuggling, there is no reason for a structure not violating a strong Freezing criterion to be delayed.[28] These data, however, are quite compatible with and predicted by UPR, assuming any account of ROE that involves, at any point during the derivation, phrasal movement from the complement of a *v*: since all phases are strong on UPR, this would constitute a violation of the Phase Impenetrability Condition (Chomsky, 2001). These data are also in conflict with Becker's claim of early acquisition for StS raising (and also undermine her claim that control verbs are analyzed as StS raising verbs, which would seem quite doubtful given these noted errors with RNE sentences). Finally, these data argue convincingly against the notion that

[28] It is certainly possible that RNE is delayed for reasons having nothing to do with UFH, but if we are to accept that the ROE results are due to UFH, and given the very minor differences between the two sentence types, this possibility seems rather doubtful. In any case, since UFH does predict that verbal passives are delayed, but not RNE sentences, under UFH there is no reason to expect a strong correlation to obtain between the acquisition of both structures. An investigation of such a correlation is at the very heart of Study 6.

young children's grammar is like that of languages that allow RNE but not ROE (e.g. Icelandic, Spanish). Unlike those languages, children here generally do not allow RNE.

Noticeably, however, the degree of non-comprehension on the RNE condition does not mirror that on the ROE condition. Many more children successfully respond to RNE sentences than ROE sentences. Fully 14 children comprehend RNE items, but fail to comprehend ROE items, while not a single child demonstrates the opposite pattern. Among the youngest half of the children tested, none comprehend ROE sentences, but 40% demonstrate above-chance comprehension of RNE sentences. Across each age group, a greater percentage of children comprehend RNE sentences than is noted for ROE sentences in Table 3.2.8. Clearly, then, some process is at work whereby results from both structures are not identical, as would be the case if UPR were correct.

In order to address this apparent difference in acquisition patterns for RNE and ROE sentences, a careful review of children's strategies for dealing with (possibly) otherwise ungrammatical structures is necessary. While non-match between levels of miscomprehension across RNE and ROE sentences is a challenge for UPR (as well as UFH, and likely any grammatical accounts), two possibilities stand out as potential explanations for these data. First, it could be that there are simply two groups of children: those for whom (something like) UPR holds (delayed on both RNE and ROE sentences) and those for whom (something like) UFH holds (delayed only on ROE sentences). The greater number of children succeeding on RNE than on ROE sentences then would simply be evidence that some children (about half) are subject to UFH, while the others are subject to UPR. Alternatively, there might be just one group of children for whom all of StS raising is ungrammatical, but two groups of children as far as compensatory strategies are concerned. That is, all children might find RNE sentences ungrammatical, but some children might make use of an interpretive strategy that happens to map to the correct responses in the current experiment, while the rest of the children are subject to some other linguistic strategy that results in non-adult (and incorrect) responses. Differentiating these two possibilities (different grammatical accounts vs. different interpretive strategies) requires a careful examination of children's response patterns and their justifications for those responses.

As already seen in Table 3.2.5, those children who do not comprehend ROE sentences, overwhelmingly demonstrate below-chance performance on this condition, indicating that they are consistently using a strategy that leads them to the incorrect answer. Extremely few children are found to be performing at chance level; children have either mastered this condition, or do not comprehend it at all in the adult manner. Going back to how the study was designed (Section 3.2.2), for the ROE condition, the scenario / test item pairings are all such that if a child applied a sort of *think*-for-*seem* RtO analysis (e.g. either the *imagine*- or *expect*-analyses from Chapter 2), he would consistently respond incorrectly. The experiment was explicitly designed to be able to test for such an interpretation. That 75% of the children failing to comprehend ROE sentences respond at statistically significant below-chance level (as opposed to the rest of the children who are just at chance; i.e. randomly guessing) is consistent with a sort of RtO *think* analysis.

Children's individual response justifications further support the idea that they are using one of the *think* RtO analyses in this condition. Consider the following two ROE test items (42)-(43) and children's response justifications (with the child's age noted in parentheses) where the child responded "true" when the correct response is "false".[29]

(42) *Ken seems to Barbie to be carrying a red backpack.*

"Cause he thinks that Barbie has a red backpack." (5;10)

(43) *Ken seems to Barbie to be in the woods.*

"Because Ken thinks she is in the woods." (4;9)

"She is in the forest, Ken thinks Barbie is in the forest." (6;1)

In all three examples, children justify their (incorrect) "true" responses by making reference to the matrix subject thinking about the other character. Likewise, in the cases below (44)-(45), where the correct response is "true" but where children have responded "false", the results are consistent with not understanding the StS raising syntax, but instead, parsing the test item with an interpretation where the matrix subject is thinking about the other character doing the action denoted by the embedded predicate.

(44) *Barbie seems to Ken to be carrying some candy.*

"Because she doesn't think there's any." (5;1)

(45) *Barbie seems to Ken to be wearing a necklace.*

"Barbie doesn't think Ken has a necklace." (6;1)

[29] Also, note that the following justifications were not "cherry picked" to make a point. Rather, these examples are just some of many, many dozen that would all lead to the same conclusion.

"Barbie was wrong, Ken doesn't have a necklace." (5;8)

Children who perform poorly on the ROE sentences hardly ever use the verb *seem* in their justifications for their responses to ROE items, where the very is used much more often in justifications for unraised items (regardless of how the child performed on the ROE sentences), and by those children who have acquired ROE. In nearly all of the cases where ROE is comprehended at below-chance level, children in fact directly use the verb *think* in the justification of their responses; the verbs *know* and *believe* are also commonly used. These children are essentially paraphrasing the test sentence, simply substituting *think* for *seem.* When faced with a ROE sentence of the form *X seems to Y to be Z,* such children tend to produce justifications of the exact form *X thinks Y is Z.* This suggests that for StS raising structures with an experiencer-phrase, children who have yet to acquire StS raising syntax consistently and incorrectly use a sort of *think*-for-*seem* analysis. While the majority of the justification responses of this sort involve the verb *think*, there are several cases of *imagine* (though none of *expect*). Again, the *imagine*-analysis is supposed to capture the semantics of *think*, just with RtO syntax. To that extent, however, it is perhaps a bit surprising that the vast majority of justification responses have a matrix experiencer verb with a finite embedded clause (e.g. all the *think* examples), though, a few examples of *imagine* + $TP_{Nonfinite}$ were recorded. While the justifications are generally consistent with the semantic entailments of the RtO *imagine*- and *expect*-analyses, the general lack of RtO verbs and infinitival complements in the justifications might also point toward a direct *seem*-as-*think* analysis on the part of the children, which would mean children are ignoring certain aspects of the surface syntax (namely the nonfiniteness of the embedded clause in the actual ROE test sentences). In either case, children's consistently below-chance responses and subsequent justifications do fit with an interpretive strategy that has children taking the matrix subject as the experiencer (thinker) and the embedded clause as the content of that experience (thought).

The RNE sentences and children's justifications for their responses are even more revealing. The fact that a majority of younger children show poor comprehension of RNE sentences is certainly evidence for a general delay in raising not linked to the presence of an experiencer-phrase. That said, many more children comprehend the RNE sentences than the ROE sentences. Turning first to those children who do not score above chance on the RNE sentences, as noted in Table 2.3.6, the great majority (75%) score below chance, once again suggesting the use of a strategy that leads to consistently wrong answers. Just as in the ROE condition, this pattern is consistent with some sort of *think*-for-*seem* analysis (i.e. the *expect*-analysis). Remember, that the scenarios and RNE test items were deliberately designed such that if a child substitutes a subject control verb for *seem* in *A seems to be Z*, he would arrive at a sentence meaning roughly A_i *thinks* PRO_i/he_i *is Z.*

Indeed, those children who answer RNE sentences incorrectly provide justification responses matching the *expect*-analysis. Consider the example in (46), where the correct response is "false", but the child answers "true". In this scenario, Barbie is not wearing a belt, but she does believe she is wearing one, and that is what the child focuses in on, and forms the basis for his justification for incorrectly accepting the RNE sentence.

(46) *Barbie seems to be wearing a belt.*

"Cause Barbie isn't wearing a belt but she thinks so." (4;9)

Consider now the examples in (47)-(48), both of which are true given the scenario, but for which children incorrectly respond "false". In both cases, children incorrectly reject the sentences as being true because the subject did not believe she was doing the action denoted by the embedded predicate, though she was. For an adult, judging the truth condition of the RNE sentences does not require one to consider the subject's thoughts, yet that is exactly what children are doing as reflected in their justifications.

(47) Barbie seems to be carrying a book.

"Because Barbie thinks she doesn't have a book." (5;1)

(48) Barbie seems to be wearing a belt.

"Because she thinks she isn't wearing a belt." (5;10)

Such focus on the thoughts of the subject character for RNE sentences, however, appears to be relegated to those children performing at a statistically significant below-chance level on this condition. The justifications for both those children at chance and above chance do not fit this analysis. Those children at chance on RNE sentences tended to give rather rambling justifications that often focused on random elements of the scenario, and which were neither germane to an adult justification nor to the subject control *expect*-analysis. This behavior is consistent with the notion that these children simply do not comprehend the structure at all and are merely guessing randomly. Those children responding at a statistically significant above-chance level on the RNE condition almost never reference the thoughts of the mentioned subject in their justifications. Further, they sometimes use the verb *seem*, which is basically never the case for those children who do not score above chance. Rather, their justifications are nearly always based on whether or not the subject was doing the action of the embedded predicate. For example, consider the RNE sentences in (49) and (50), where in both cases the correct answer is "false" and the children respond correctly.

(49) *Barbie seems to be wearing a belt.*

"Because she doesn't have a belt." (4;11)

(50) Barbie seems to be carrying a book.

"Barbie doesn't have a book in her bag because Ken took it." (7;6)

As for their justifications, both children focus on whether or not the subject is actually doing the action denoted by the embedded predicate. Such responses are consistent with either adult StS raising or the copula-analysis, under which the child is hypothesized to replace the raising verb (here *seem*) with the copula. Interestingly, the first justification (49) is given by a child who failed to comprehend ROE, while the very similar justification in (50) is given by a child who does comprehend ROE. As mentioned in the experimental design section above, while this study can differentiate acquisition of RNE syntax from lack of RNE syntax plus the *expect*-analysis, it cannot differentiate the acquisition of RNE syntax from the lack of RNE syntax plus the copula-analysis, since in the latter case, both map to the same truth conditions, but by way of different syntactic analyses. This is very important because it means that while failure on the RNE condition is quite telling concerning children's (lack of) knowledge of StS raising, success on the RNE condition is ambiguous.[30]

If younger children, most of whom are found to not comprehend ROE sentences, in fact lack StS raising syntax altogether (not just ROE), then apparent success on RNE sentences might not be support for UFH, but rather, is equally consistent with UPR and the copula-analysis. That is, all children might lack the necessary syntax to compute RNE sentences, but while some adopt the *expect*-analysis, other children instead adopt the copula-analysis, which just happens to lead to adult responses on the RNE condition. While the *expect*-analysis affords a generally consistent analysis (though by means of different syntactic parses; i.e. RtO vs. subject control) by which to interpret both ROE and RNE sentences, the copula-analysis might actually be a more helpful heuristic.[31] In

[30] A nice way to look for evidence of the copula-analysis using the already collected data, which will have to await future motivation and effort, would involve the following two approaches. First, one could investigate whether younger and older children give different types of justifications for their correct responses to RNE sentences. Second, one could look for differences in how children who comprehend ROE sentences justify their correct responses to RNE sentences versus those children who have yet to comprehend ROE sentences. In both cases, there is an expectation that there will be a detectable difference between justification due to a truly adult StS raising interpretation and one involving the copula-analysis. Alternatively, one could simply construct an experimental paradigm which puts StS raising and copula interpretation into conflict. Such was the attempt in Becker's second experiment, and forms the basis for Study 5. Finally, one could investigate children's interpretation of StS raising sentences with inanimate subjects. These sentences are expected to be incompatible with subject control interpretive strategies, and might lead to a chance, as opposed to a below-chance, response pattern among premature children. While such an investigation is not included in this dissertation (though subject animacy is explored in child natural production data in Study 7), it could prove important for those interested in details concerning compensatory heuristics. Ultimately, the current work focuses on whether children find StS raising grammatical, not the myriad potential interpretations children might give to otherwise ungrammatical inputs.

[31] The copula analysis, however, is really only applicable with RNE sentences. There is no obvious way to accommodate the experiencer-phrase in ROE sentences on the copula-analysis:

(i) Barbie is *(to Ken) wearing a scarf.

This point is taken up in Section 3.3.5.

particular, only the copula-analysis is consistent with inanimate subjects, since the subject control *expect*-analysis for RNE sentences requires animate (sentient/intentional) subjects. The possibility of the application of the copula-analysis by some of the children included in this study requires further investigation.

3.2.6 *Conclusion*

The experiment presented here assessed children's comprehension of two types of StS raising sentences: those with an experiencer-phrase (ROE) and those without one (RNE). The gathered data suggest that children do not comprehend StS raising either over or with no experiencer, though unraised sentences with an experiencer are comprehended. These data are incompatible with the tested English-speaking children believing, for example, English to be like Spanish (see Section 3.1.2), which could capture the difficulties with ROE sentences, but would offer no explanation for problems with RNE sentences. It would thus appear that children's difficulties with StS raising are not due to any form of parameter missetting as relates to whether or not a specific language licenses ROE structures. Of the two maturational theories under serious consideration, it is UPR that better captures this behavioral pattern, and not UFH. This behavior is in line with the predictions of UPR, which asserts that *all* subject-to-subject raising structures are ungrammatical for immature children. Furthermore, we find no evidence for Becker's claim that children comprehend StS raising from an early age. Strong evidence is found of a single *think*-for-*seem* analysis among those children who do not comprehend ROE sentences (e.g. the RtO *imagine*- and *expect*-analyses). When presented with RNE sentences for which their grammar offers no adult interpretation, however, children have (at least) two options. While many children utilize a (reflexive) form of some *think*-for-*seem* analysis (e.g. the subject control *expect*-analysis), other children ignore the matrix verb and treat the sentence as a copular construction. Only in the former case are problems with RNE comprehension detectable in this experiment, as use of the latter (copula-)strategy is indistinguishable from adult interpretation (given this particular experimental paradigm, not, thankfully, in principle). How then to differentiate these two possibilities?

If, as is predicted by UPR, *all* subject-to-subject raising is delayed for children for the same (maturational) reason as results in verbal passives being delayed in development, one would expect to find that children who comprehend verbal passives will comprehend RNE sentences, and vice versa. Any strong positive correlations between comprehension of verbal passives and RNE would of course be unexpected on UFH. These predictions will be taken up in Study 6. First, however, let us turn to an experiment that attempts directly to distinguish the application of the copula-analysis from true knowledge of StS raising in RNE sentences.

3.3 Study 5: Raising with no Experiencer and Subject Control

3.3.1 *Motivations*

Study 4 offers further evidence that unraised sentences with an experiencer-phrase are generally acquired at an early age, and significantly sooner than ROE sentences that have the same meaning. The acquisition of RNE sentences appears to also be generally delayed, at least for many young children. Yet, the pattern of acquisition for RNE does not exactly match that for ROE, at least in the previous study. As was observed in Study 4, every child who does not comprehend ROE sentences also fails to comprehend RNE sentences, while the opposite does not hold. There are numerous cases of children successfully responding to RNE sentences, but who fail with ROE sentences. These data are compatible with the idea that while nearly all young children find ROE ungrammatical, some find RNE ungrammatical while for others RNE is grammatical; that is, two groups of children might exist based on different maturational constraints on their grammar, some whose grammar is governed by (something like) UPR (both ROE and RNE are ungrammatical) and others whose grammar is governed by (something like) UFH (only ROE is ungrammatical).

As noted in the previous section, however, Study 4 is only designed to detect those errors on RNE sentences that are compatible with the application of the *expect*-analysis (i.e. a reflexive *think* interpretation). Indeed, any child adopting a copula-analysis for RNE sentences would actually respond correctly to the RNE test items in Study 4. This raises the alternative possibility that the hitherto observed mixed performance on RNE sentences is not due to two different grammatical acquisition deficits, but only to a single deficit on which all StS raising is ungrammatical, whereby two groups of children exist on the basis of which strategy they adopt for dealing with what would otherwise be an ungrammatical structure. On this idea, UPR might hold of all premature children, with some hitting upon the *expect*-analysis and some on the copula-analysis. Certainly the simpler hypothesis is that which posits a single group of children (UPR) as opposed to two groups (UPR, UFH) to account for the mixed data in Study 4.

What then to make of Becker's (2005, 2006) data, which apparently demonstrate successful (i.e. adult) interpretation of RNE sentences for nearly all young children, seemingly at odds with the findings reported in Study 4, and on whose standing Becker claims StS raising is acquired early in linguistic development? Becker offers three pieces of evidence that even young children have acquired StS raising: (1) successful grammaticality judgments of RNE sentences in her first experiment, (2) grammaticality judgments of sentences with subject control verbs in her first experiment that are compatible with StS raising interpretation, and (3) successful truth-value judgments of RNE sentences in her second experiment. The question then arises of how to reconcile Becker's data and claims with our data demonstrating great problems on ROE (unexpected if StS raising is generally fine), and specifically, what to make of the many children who failed to comprehend RNE sentences in Study 4.

Note that the raising finding from Becker's first study is actually quite compatible with children's early grammar lacking StS raising and the application of the copula-analysis. Her first experiment involved having children judge the grammaticality of sentences consisting of an inanimate subject, a matrix StS raising or subject control verb, and an embedded infinitival clause whose predicate was either compatible or incompatible with an inanimate subject. Becker finds that children as young as three-years-old tend to respond correctly, accepting raising sentences with compatible predicates (51) and rejecting raising sentences with incompatible predicates (52). Observe, however, that the application of the copula-analysis would result in exactly this behavior as well. A child asked to judge the grammaticality of (53) and (54) would also likely respond that the sentence with the compatible predicate is fine, but that the one with the incompatible predicate is problematic.

(51) The hay seems to be on the ground.

(52) #The hay seems to be excited.

(53) The hay is on the ground.

(54) *The hay is excited.

Since the subject is inanimate, there is no hope of the subject-control *expect*-analysis applying (55), accounting for the generally uniform performance across all children. As such, the raising data from Becker's first experiment cannot be taken as strong support for early acquisition of StS raising, as it is just as compatible with StS raising being delayed and the use of the copula-analysis, which appears independently necessary to account for the data from Study 2.

(55) *The hay expects to be excited.

Becker, however, does not only rely on children's successful responses to the two raising conditions from her first experiment from which to argue that StS is early in development. She points out that young children's failure to provide adult grammaticality judgments for control verbs also offers evidence of children's grammar having access to StS raising operations. When asked to judge the grammaticality of (56) and (57), both of which adults reject (control verbs require sentient/intentional subjects), the three- and four-year-old children Becker tested did correctly reject the control cases with incompatible predicates, but they incorrectly accepted those with compatible predicates. Becker takes such a dichotomy to indicate that these younger children lack subject control, and instead interpret the control verbs as if they were non-thematic raising verbs (see Section 3.1.1). As such, children would basically parse (56) and (57) as if they meant (58) and (59), and therefore, since raising verbs themselves place no thematic restrictions on their matrix subject, children would only reject those control sentences with incompatible predicates (which is indeed how they respond). Becker's second experiment, discussed briefly in Section 3.1.1 and to which we return in detail below, offers further evidence suggesting StS raising is delayed.

(56) #The flower wants to be pink.

(57) #The flower wants to walk away.

(58) The flower seems to be pink.

(59) *The flower seems to be excited.

While we acknowledge the ingenuity of Becker's experiments (especially the second, which really is quite clever), there are a number of important problems with them, both experimental and conceptual, which ultimately cast serious doubt upon her claims about StS raising and subject control, and to which we will now turn. Most doubtful is her claim concerning subject control verbs being non-thematic StS raising verbs. For one thing, there is no shortage of experimental evidence (reviewed shortly) and commonsense/anecdotal evidence that children *do* correctly understand subject control (as a grammatical process) and subject control verbs (as being different from other verbs).

While UPR and UFH (as well as ACDH and EARH) predict that certain structures involving long-distance dependencies will be quite delayed for children (e.g. verbal passives and ROE), an important consideration is that not all structures involving non-adjacent dependencies will be delayed nearly the same way, for example, see Section 1.2.3 for evidence that *A'*-movement is acquired early. There is another class of long-distance structures, though, that is also important to consider, because these structures seem to involve relations between argument positions: this is the class of control structures (see Section 1.3.1). In a sentence like (60), PRO is "controlled" by the subject *John*. Let us call these sentences like (60), which have the same form as those tested in Becker's two experiments, cases of "obligatory control" (OC), since PRO cannot be interpreted as referring to some non-mentioned (external and non-c-commanding) entity.

(60) John tried [PRO to leave]

The standard analysis of control does not take the relation between *John* and PRO in (60) to be an Agree relation (much less a Move relation; cf. Section 4.1.2). Rather *John* and PRO are co-indexed or made co-referential (or PRO is made to be referentially dependent on *John* in some other way, perhaps in the semantics). Thus UPR and UFH do not predict any difficulties with OC cases like (60).[32] Nevertheless, in some structural respects, the relation between the controller and PRO in a sentence like (60) seems

[32] EARH also does not predict any difficulties for (60) since even the embedded sentence has an external argument, PRO, which has referential content (it is not an expletive).

similar to an Agree or Move relation; the controller c-commands and is fairly local to the controlled element. Similarly a moved element c-commands and is fairly local to the position from which it moved (Sections 1.2.1 and 3.1.2). The same holds generally for Agree. The major difference seems to be that the control relation is not sensitive to phases in the way that Agree and Move are. There is a non-defective, phasal *v* in (60), the *v* that selects the VP *tried PRO to leave*. PRO is in the complement of this phasal *v*, yet there is no problem in relating it to its controller. Phases do not seem to play the same type of role in control as they do in Agree or Move. Thus, UPR does not predict a problem for control, and since no Smuggling seems necessary on any standard accounts of control, UFH also predicts no delay. Of course, the relevant question is one of empirical verifiability: does control develop earlier than structures that depend on non-phasal (weak) *v* or exceptions to Freezing? If so, this would be further evidence for UPR or UFH. Alternatively, if OC develops as late as verbal passives and ROE, then one would have evidence against UPR and UFH, and in favor of a problem with all relations that appear to involve local c-command.[33]

Reviews of the development of control can be found in Wexler (1992) and Guasti (2002). McDaniel, Cairns, and Hsu (1990) argue on the basis of their experiments that, "there is a stage, previously unattested as far as we know, during which children lack control." Wexler (1992), surveying the data, concludes, "there is an early stage in which children don't know that the empty subject in complements and adjuncts must be

[33] Of course, such local c-command structures would include Principle A of the binding theory, and it is known that reflexive binding develops much earlier than passive structures, roughly around age three, depending on the quantitative standards and types of experiment used (Wexler and Chien, 1985, Chien and Wexler, 1990; among many others). Such phenomena already suggest that UPR or UFH is more on the right track than difficulties with local c-command relationships. Of course, a similar timeline of (late) development of OC, verbal passives, and StS raising might also suggest that UPR is on the right track and that a movement analysis of control is warranted (see Hornstein, 1999; and Section 4.1.2).

controlled." This holds, however, only at very young ages. Wexler notes that of 20 children from 3;9 to 5;4 years-old, only one of the four youngest children (3;9-3;10) lacked complement control (OC) and only one of the 16 older children (3;11-5;4) lacked OC.[34]

Sherman (1983) conducted a comprehension study using an act-out task with sentences like *Mary told John PRO to leave.* The group from 5;0 to 5;11 had a mean accuracy of 81%. These numbers are much better than what we see on ROE or RNE, where in the StS raising Studies 1-4, groups around this age respond at no better than chance level. Guasti (2002) reviews many studies and agrees that initially children lack control in certain complement constructions, but "by 3 years children know that PRO is distinct from lexical pronouns."

The consensus seems to be that OC is in place at about three years of age. Before this, external control is possible for children. This is a much earlier age than the six- to seven-year-old age range in which StS raising appears to become mastered in Studies 1-4. Wexler (1992) proposed that until about age three, children over-extend the case filter, due to a maturational process, requiring that all NPs have Case. This would outlaw PRO. Whatever the explanation, control of complements is mastered much earlier than StS raising or verbal passives of subject experiencer verbs in English.[35]

Of course, children also use control verbs (like *try* and *want*) quite often. Pinker (1984) argues that in natural production, the use of an external controller for PRO is almost non-existent in the data that he analyzed.[36] Our own cursory search of the

[34] We are ignoring adjunct control, which complicates the picture somewhat, but for different reasons. See Wexler (1992).
[35] We return to the question of the very small set of subject control verbs like *promise* in section 4.2.
[36] Pinker's observations are written in terms of Equi, but the data are equivalent to what we note.

CHILDES corpus of children's utterances (MacWhinney, 2000) turns up thousands of *want* control constructions, such as the following examples in (61):

(61) [I want to read this paper]

[I want to shave too]

[I want to have some espresso]

[um # I want to be a gypsy]

[I want to drive]

[I want to ride on a panda]

[I want to hold a lamb # I didn't hold it]

Needless to say, it is not very plausible that the three-year-old speaker of the third utterance in (61) above means to convey that he actually *seems* (Becker's control-as-raising idea) to be having an espresso (the full transcript makes it clear that he does not). By the same token, the other utterances are odd indeed on the assumption that the children are using *want* to mean *seem* (or any other relatively semantically vacuous StS raising verb, as required on Becker's hypothesis).

Corpus-based evidence for children's correct interpretation of control verbs is not limited to production. Even when responding to a parent, young children show an unambiguous understanding of verbs like *want*, as in the following example between a mother (MOT) and her child (CHI) in (62), just one of many hundreds of such cases. It is unclear what to make of this exchange if we are operating under the assumption that children interpret control verbs as raising verbs.

(62) MOT: do you want to do that again?

CHI: (o)k.

(bates/free20/hank20.cha:240)

Also in regard to Becker's claim that children lack control, there exist striking contradictions from Becker's own data. Becker reports that her subjects performed well on the control verb condition in her second experiment, yet, according to her hypothesis children should be interpreting these verbs as StS raising verbs, which should have produced wrong answers. According to Becker's hypothesis, children interpret *want* as *seem* (or some other semantically-neutral raising verb). Thus, if they are indeed parsing the main verb, they should interpret the test sentence *The pig wanted to eat the donut*, which given her scenario is true, as *The pig seemed to eat the donut*, which given her scenario is false, and should prompt the children to respond incorrectly as such. Children, however, performed very well on the subject control condition in the second experiment. On the hypothesis that children interpret control verbs as StS raising verbs, the results of Becker's second experiment are unexplained. On the commonsensical hypothesis that children readily comprehend these control verbs, these results are unproblematic.

Furthermore, Becker's idea that children treat control verbs as raising verbs makes (at least) two syntactic predictions against which extensive production data speak. First, if control verbs like *want* are actually raising verbs, then they should not allow bare DP complements. While such control structures are allowed in the adult grammar (e.g. *The man wants an apple*), they should be ruled out for the child since bare DPs are not grammatical with raising verbs (e.g. **The man seems an apple*). Yet a brief glance at data on the CHILDES corpus turns up thousands of examples of control verbs with bare DPs. Second, if control verbs are interpreted as raising verbs, then a control verb like *want* might very well be expected to have an "unraised" counterpart (e.g. *It wants that the flower is pink*). Yet, there is no evidence from production data that children ever use control verbs in such a manner (with expletive subject and finite complement). Children's use of bare DPs with control verbs, children's lack of "unraised" forms with control verbs, and Becker's own second experiment, plus twenty years of research on the acquisition of control strongly speak against Becker's claim that children provide a StS raising analysis to sentences containing subject control verbs.

How then to explain Becker's finding that certain young, three and four year-old children accept control verbs with inanimate subjects and compatible embedded predicates? One plausible alternative explanation for this finding that children accept sentences like *The flower wants to be pink* as grammatical is simply that they operate under the assumption that the sort of cartoon inanimate characters used in the experiment are intentional beings (either for the purposes of the experiment or more generally). This is particularly likely in a story-based, game-like experimental setting, where children are often willing (and even encouraged) to suspend normal judgments and anthropomorphize pictured objects.

Becker (2005) does not consider this possibility that children might extend intentionality to her intended non-intentional (inanimate) subjects. The possibility is raised, but dismissed in Becker (2006). In her Footnote 3, she argues against it on the basis that (1) children rejected the non-compatible predicates (e.g. *be hungry, be excited,* *be friendly*) for the control condition, and (2) children sometimes gave justifications consistent with distinguishing between animate and inanimate subjects (e.g. "Flowers aren't alive"). The first point is quite valid. If children *are* extending intentionality to the cartoon characters, they must only be doing so for "inherent" properties (i.e. those expressed in the compatible condition; e.g. *be pink*, *be small*).[37] In any case, this is also an issue for Becker, as her raising-for-control hypothesis actually predicts that young children should be at below -hance level on their grammaticality judgments for the control-compatible condition, not the observed chance level. This point goes ignored, and without data breakdown by either item or individual child, it is hard to address, especially since, and this is quite relevant, the experiment only tested two items per condition, which is quite minimal.

As for Becker's second claim that children's justifications support an awareness of animacy distinctions, this is simply not supported by her own data. The fact that the subject control results of the first experiment *are* due to younger children taking the inanimate subjects to be sentient in this experiment is strongly suggested by children's justifications for accepting the control sentences with inanimate subjects (Becker, 2004), two examples of which are given in (63)-(64).[38]

[37] There is also the possibility that children have simply applied the copula-analysis to the control condition in Becker's first experiment (as is being suggested for the raising condition). None of the pictures Becker uses depict the inanimate subject doing the "action" denoted by the embedded predicate in the incompatible condition, but they *do* depict the "action" of the embedded predicate in the compatible case (M. Becker, p.c.). So for example, the picture associated with the sentence *The flower wants to be pink* actually shows a pink flower, so any children using the copula-analysis would arrive at *The flower is pink*, which is perfectly acceptable, though the actual control sentence is not. On this account, however, one would still have to explain chance, as opposed to below-chance, performance (possibly due to only some children applying the copula-analysis), and why children should ignore the control verb, especially given the discussion covering the extensive evidence that children at this age comprehend subject control.

[38] Note, that these examples were not included in either Becker (2005) or Becker (2006), but come from a handout Becker prepared concerning the experimental data presented in those two papers.

(63) Test item: # The bucket wants to be in the sandbox

Child: I think the bucket should be in the sandbox.

Investigator: But do you think the bucket could want to be in the sandbox?

Child: I think so. (age 3;11)

(64) Test item: # The flower wants to be pink

Child: And the bees want to eat them!

Investigator: Do you think the flower could want to be pink?

Child: Yes, and green too! (age 3;1)

Given the strong asymmetry between the age of acquisition for subject control versus verbal passives and ROE (and perhaps RNE), UPR and UFH are supported since they predict this asymmetry, while Becker does not. The data from Becker's first experiment instead appear to be explained by young children allowing subject control verbs with cartoon inanimate subjects (taken to be intentional) and embedded predicates expressing inherent properties, and by children applying the copula-analysis to RNE sentences.

Becker's second study, however, is explicitly designed to rule out an interpretive strategy like the copula-analysis in which children are essentially ignoring the matrix (raising) verb. This experiment, as relates to the StS raising condition, involves stories that attempt to distinguish appearance from reality. So, in one case, a white dog walks under a black light, thus appearing purple, and children are asked to judge the truth of *The dog seemed to be purple*. Becker's idea is that if children ignored the matrix verb they would derive *The dog ... purple*, and since the dog was truly white, should incorrectly reject it. Children successfully respond to the sentences of her second experiment, leading Becker to conclude that StS raising is in place at a very young age.

Again, the ingenuity of this second experiment cannot be overstated, especially since it provides a very natural way to test for the grammaticality of StS raising without having to rely on an experiencer-phrase (as in Studies 1-2) or on constraints on detectable interpretive strategies (as in Study 4). That said, there are certain issues with how the study was conducted that raise serious questions about the validity of the gathered data. The first such concern revolves around how to reconcile Becker's claim that children as young as three years-old comprehend StS raising sentences with the finding from Study 1 that most three-year-olds (60%) fail to comprehend the unraised counterparts of those very raised sentences. If these young children do not even comprehend unraised sentences, one wonders how it is that they can comprehend the raised equivalents. It is unfortunate that Becker failed to include unraised sentences, as it means that one cannot discern which children (and how many) did not comprehend the raising verbs to begin with. This is an important oversight, as it is unclear what to conclude about a child's knowledge of StS raising if he does not understand the meaning of the StS verb used in the test. Again, when 37% of three-to-five year-old children in Study 1—the age range Becker also examines—do not comprehending the unraised sentences, this suggests that many children do not even know the raising verbs, and one should be cautious of claims that children somehow nevertheless comprehend StS raising.

An alternative explanation exists for Becker's data, however, which is compatible with the findings from Studies 1-4: namely, that her findings for StS raising are not valid.

This has to do with how Becker attempted to experimentally control for the possibility that children might ignore the matrix StS raising verb if they could not interpret it in the relevant structure. She writes of her second experiment, "A child parsing only *the dog ... be purple* should respond "false", since the dog was not in fact purple; but a child parsing *the dog seemed to be purple* should respond "true" since the dog did seem to be purple when standing under the lamp." Though children indeed respond correctly to the test items in this second experiment, Becker's test items might be confounded. Even on a copula-analysis, children might respond in line with Becker's findings if they confuse how the copula sentences relate to the scenarios, as described below.

We conducted an informal investigation into how native English-speaking adults respond to the test sentences in Becker's second experiment. When talking with our informants, it became clear that two mechanisms in particular might account for children's successful responses to RNE sentences even if those children lack StS raising and instead make use of a copula-analysis, which is precisely the analysis Becker had set out to avoid.

Consider the case of the white dog walking under the black (i.e. purple) light and asked to judge the truth of *The dog seemed to be purple.* Becker tells us that if the child cannot interpret this sentence (e.g. because he lacks StS raising, perhaps due to UPR), he might attempt to ignore the matrix verb, arriving at *The dog ... purple.* This is her example of the copula-analysis, but notably, Becker's analysis does not explicitly address the realization of the copula itself (i.e. her use of "..."). A child not comprehending the matrix verb (here *seemed*) is still very likely to notice that it contains past tense morphology. Thus, if a child did recognize *seemed* as a past tense form, but did not know its meaning, he would most likely substitute the past tense form of the copula (i.e. *was* not *is*), and his parse would be *The dog was purple.*

It is here that the data from our adult subjects proves quite relevant as it provides great insight into possible reasons for why children might have responded as they did in Becker's second experiment. When we asked the adults to judge the truth of this copula sentence, with past tense morphology, given the described context, many reported the sentence to be fine ("true"). They justified their responses by stating that the dog was purple under the light since his fur looked purple. If asked to judge instead the same sentence, but with a present tense copula, *The dog is purple*, most adults then responded oppositely, saying this was "false". For these adults, there is a difference between how the past and present tense copula relates to truth-value judgments of appearance versus reality. The past tense copula *was* is compatible with mere appearance, while the present tense *is* requires reality. The past tense *was* implies a change of state, while *is* reflects inherent properties. Since the test subjects are free to interpret the sentence with respect to any moment in the story scenario, the fact that there *does* exist a moment when the dog appears white and another moment when he appears purple, the use of *was* is more felicitous than *is* (with reference to the adjectival predicate *purple*).

Also of interest, when we asked adults to judge the sentence *The dog was white* given the same context as above, most accepted this to be true, but when asked for their justification, all referenced the moment before the dog walked under the light in the scenario. Even though the dog is truly white at all moments of the scenario, the adults justified their response by referencing the moment before the dog walked under the purple light. This is different than when adults were asked to judge *The dog was purple,* in which case justifications all made reference to the moment the dog is standing under the purple light. It would seem then that when the past tense form of the copula is used, (adult) responses are based on referencing different time points during the scenario. This is an added complication, since no such effect appears relevant for judging StS raising sentences or subject control sentences. If children are using a copula-analysis when interpreting RNE sentences, the effect of tense must be addressed.

Crucially, and of relevance below, a similar dichotomy in adult judgments was also noted when those adults accepting *The dog was purple* were pressed if the dog *really* was purple. All answered "no". Thus, it is clear that adults can make such appearance-reality distinctions, but that to do so, present tense morphology or the explicit inclusion of the adverb *really* makes a big difference. This is of course very relevant, since Becker's entire premise for rejecting the copula-analysis as an explanation for her StS raising results is that ignoring the matrix verb should produce errors. Yet, since Becker never explicitly tested sentences without the raising verbs (i.e. copula sentences), it is not at all clear that children might not respond correctly to RNE sentences, but only on the basis of a copula-analysis that includes tense confusion or failure to differentiate reality and appearance. By not including a condition that directly tests how children respond to a copula sentence (especially one that has the same tense marking as the subject control and StS raising sentences), Becker never establishes that in her test scenarios children are actually distinguishing what truly appeared to be the case from what really was the case.

Indeed, we attempted a replication of Becker's second experiment, but also included a copula (experimental) control condition as well.[39] These copula

[39] The experimental details of our attempted replication are basically those described in 3.3.2 (i.e. the experiment in Study 5).

sentences contained the past tense form of the copula (i.e. *was*), to match the past tense use of the raising and control verbs. We replicate Becker's findings that children successfully respond to both subject control and StS raising (RNE) sentences. Crucially, however, nearly every child failed to respond "correctly" to the newly included copula sentences. From their justifications, it was clear they were falling prey to the same issues that plagued our adult test subjects, namely, they were "incorrectly" basing their judgments on appearance, and not reality. It is therefore clear, on the basis of these experimental data, that one would not yet want to conclude that children comprehend StS raising since they failed to comprehend the copula sentences. Importantly, in follow up pilot work, it was also clear that children *can* give "correct" judgments to the copula sentences. This was accomplished by changing the copula to its present tense form (i.e. *is*) and including the adverb *really* (e.g. *The dog really is purple*). What all this pilot work suggests is that meaningful conclusions concerning children's RNE sentence comprehension using (something like) Becker's second experiment can only be made once children have explicitly been shown to comprehend copula sentences, otherwise, the copula-analysis cannot be ruled out. Becker's intention (presumably) in her second experiment had been a task that can differentiate **at the time the dog is standing under the purple light** between appearance (i.e. *the dog seems to be *white/purple*) and reality (i.e. *the dog is white/*purple*). However, by conflating issues of tense and strictly differentiating appearance and reality, it appears she was unsuccessful in doing so.

While a few studies have examined children's knowledge of RNE to date (namely, Becker's two experiments and Study 4), further experimentation is very much needed, especially as relates to ruling out the copula-analysis as detailed above.

3.3.2 *Experimental Design*

The basic idea behind the experimental paradigm to be employed in the current study is to replicate Becker's second experiment, but to ensure that children are familiar with the StS raising verbs being tested (by including unraised sentences with these verbs) and to ensure that children are in fact sensitive to the distinction between reality and appearance (by including a copula condition, making use of picture-tense pairings, and including the adverb *really*). Doing so will allow more faithful interpretation of results concerning children's knowledge of subject control and RNE.

As already discussed, Becker's second experiment relied crucially on children differentiating cases of reality from appearance. Unfortunately, she did not include any experimental controls to verify that the children she tested in fact were sensitive to this experimental manipulation. Several changes will therefore be introduced to help ensure children are responsive to this contrast.

The methodology to be used to assess children's comprehension of the relevant structures is a Truth-Value Judgment (TVJ) task (Crain and Nakayama, 1987, Crain and Fodor, 1993). This paradigm involves the child participating in a "game" in which he and a puppet (Mr. Bear) observe and hear about a sequence of pictorial representations of a scenario. The child is told that Mr. Bear will offer some comment at the end of each scenario about what occurred, but that he is somewhat silly and will often make mistakes.

The child is told that he must serve as Mr. Bear's teacher for the day, and let him know when he is right or wrong, and why.

In this study, scenarios consist of dialogue and two pictures, *initially* presented together, side-by-side, ordered temporally with respect to the story told by the experimenter. The stories and pictures are used to introduce a difference in reality versus appearance. The scenarios employed are similar, or in some cases, identical to those used in Becker's second experiment. As a concrete example, consider the two pictures representing a single scenario, depicted below, in Figure 3.3.1. In the first picture, a white dog is standing next to a purple light, and remarks that he does not want to go under the light, because it will make people think his fur is purple, a color he finds very ugly (unlike his white fur which he finds to be very beautiful). In the second picture, he has nonetheless gone under the light, making his fur look purple, although a tiny portion of his tail remains outside of the light, to remind the child that in reality, his fur is white.[40]

***Reality*: The dog IS white.**

---

[40] The other three scenarios involve [picture (i) and (ii)]:

(1) (i) a pig that cannot swim standing on a river bank, (ii) looking into the river so all you see is his reflection in the water (as if he were in the water).

(2) (i) a very small horse who very much enjoys the benefits of being small, (ii) but who appears quite large under a magnifying glass.

(3) (i) an elephant who greatly prefers the warm sun to the cold shade, and who is clearly standing around during the sunny day time, (ii) but who looks to be in the shade when the puppet puts on some sunglasses.

Appearance: **The dog LOOKS LIKE he is purple.**

Figure 3.3.1: Pictures depicting a single scenario that are used to test all four conditions in Study 5. Two pictures help differentiate reality from differing appearance. At the time the test sentence is read, only the second picture remains in view of the child.

Following every scenario, the child is asked three comprehension questions to establish the facts of the story before Mr. Bear comments on what has occurred; with respect to the scenario in Figure 3.3.1, these questions consist of: "What color is the dog in real life?", "What color does the dog look like he is?", and "What color does the dog want to be?"[41] These questions serve to establish that the child understands the details of the scenario, such that any noted errors cannot be attributed to the child being incorrect/confused about the pertinent information in each scenario. If the child does not answer all of these pre-test questions correctly, the scenario is simply repeated.

Once the child has demonstrated that he understands the scenario, the first picture (e.g. The dog standing next to the light) is removed, and Mr. Bear comments *only* on the second picture. Mr. Bear comments using one of four sentence types: copula sentences

[41] It is interesting to note that all children who correctly answer the final pre-item question have shown evidence of correctly comprehending at least one obligatory control verb, which Becker (2005, 2006) claims young children should fail to do.

(65), subject control sentences (66), unraised sentences without an experiencer (67), and RNE sentences (68). Following Becker (2005, 2006), *seem* and *appear* are used as the two StS raising verbs, and *like* and *hate* as the two subject control verbs. All verbs appear in their present tense forms. The truth of each sentence is manipulated by the choice of embedded predicate.[42]

(65) The dog is really white (T) / purple (F).

(66) The dog really likes / hates to be white (T/F) / purple (F/T).

(67) It really seems / appears that the dog is white (F) / purple (T).

(68) The dog really seems / appears to be white (F) / purple (T).

The copula condition (65) is crucial for determining that children are indeed sensitive to reality even when given conflicting appearance. That is, successful responses to the copula sentences establish that children understand that mere appearance is not reality. These sentences also serve as general experimental control items, requiring the child to pay attention to somewhat complex scenarios, to reward and correct a test puppet, and to justify his responses. Answering the copula sentences correctly will be a prerequisite for inclusion in subsequent analyses, since it makes little sense to ask if children can comprehend the unraised and RNE sentences, which both hinge on appearance, if children cannot determine what actually holds in each scenario.

---

[42] The subject control condition is complicated by the fact that the truth-condition of the sentence is manipulated by the choice of embedded predicate (as is the case in the other three conditions) and the choice of the particular control verb (not an issue in the other three conditions). This is explicitly done in order to help distinguish between competing theories concerning the acquisition of control discussed below.

The subject control sentences (66) are included for a number of reasons. In order to respond correctly to these sentences, children must comprehend the details of the various scenarios. Furthermore, children must have no general difficulty comprehending bi-clausal sentences. Crucially, of course, these sentences require children to actually comprehend syntactic subject control. The sentences and scenarios have been specifically crafted to detect possible errors due to children conceivably misinterpreting subject control in several ways.

Becker's claim is that children treat subject control verbs as semantically vacuous StS raising verbs. If children indeed treat both *like* and *hate* as (something like) *seem*, they should respond in the same manner regardless of which particular control verb is used. Remember, however, that the scenarios are all designed so that in order to answer correctly, different responses are required for each subject control verb. As such, any child treating the two control verbs as a semantically vacuous raising verb will be at chance performance on the condition, responding correctly for one control verb and embedded clause but incorrectly for the other, and vice versa depending on the embedded clause. Which verb is answered correctly is dependent on the different scenarios and particular embedded clause. Consider the case of the white dog who finds the color purple ugly, but thinks his white fur beautiful. If Becker is correct, asked to judge *The dog likes to be purple*, children will interpret this as akin to *The dog seems to be purple* and (since according to her, children comprehend StS raising syntax) respond (incorrectly) "true". Asked to judge *The dog hates to be purple*, children would also analyze this as *The dog seems to be purple* and (now correctly) respond "true". When the embedded predicate *to be purple* is changed to *to be white*, the opposite pattern should be detected. Therefore, Becker's hypothesis about children treating subject control verbs as StS raising verbs can be confirmed by chance performance on this condition and the particular pattern of responses given different verb and embedded clause pairings.

An alternative misanalysis is that indeed children do not comprehend subject control, but rather than applying Becker's "control-as-raising"-analysis, children apply a copula-analysis to the subject control sentences, replacing the control verb with the copula. Once again, children would be collapsing *like* and *hate* to a single interpretation, but crucially, the copula-analysis makes the opposite predictions as Becker's hypothesis. Asked to judge *The dog likes to be white* on the copula-analysis, children would interpret this to mean *The dog is white*, and would (correctly) respond "true". When asked to judge *The dog hates to be white*, if children take this to also mean *The dog is white*, they will again respond "true", but this time will have answered the actual test sentence incorrectly. The opposite pattern should be detected for the embedded clause *to be white.* The copula-analysis can therefore be detected on the basis of chance performance, with a particular response pattern opposite that of Becker's hypothesis.

If, however, children do comprehend subject control, as suggested by past research (see Section 3.3.1), they should respond correctly to all subject control sentences, regardless of the different control verb-embedded clause pairings. By employing two subject control verbs with semantically opposite meaning, various possible interpretive analyses for the control sentences can be differentiated.

The inclusion of unraised sentences (67) serves a crucial role that was missing in Becker's own study. Successful comprehension of the unraised sentences establishes that children are sensitive to the experimental manipulation of appearance, without conflating the issue of the acquisition of syntactic raising. In order to respond correctly ("true") to *It seems that the dog is purple*, a child must differentiate between the dog appearing purple under the ultraviolet light and the fact the dog is actually white. If children ignore the matrix StS raising verb, and only parse the embedded clause, they will consistently respond incorrectly to the unraised condition. That is, interpreting *It seems that the dog is purple* as *The dog is purple* will lead to below-chance performance. The verbs *seem* and *appear* both serve to denote appearance as different from reality. By including both copula sentences (reality) and unraised sentences (appearance), awareness of this crucial experimental manipulation (appearance vs. reality) can be confirmed on an individual subject basis prior to ever asking if a particular child comprehends StS raising.

Including unraised sentences is also important given that many children in Study 1 did not comprehend such sentences, and this alone raises troubling questions for Becker's claims that young children have acquired StS raising syntax. The unraised sentences tested here will not include an experiencer-phrase (as was the case in Studies 1-4) in order to serve as more faithful controls for the RNE sentences, which of course do not contain an experiencer-phrase. Unlike in Study 4, where RNE sentences were also tested, the presence of the experiencer is not necessary for successful response in this study. By determining which children comprehend unraised sentences with *seem* and *appear*, one can establish which children comprehend the lexical meaning of the raising verbs independent of StS raising. Since unraised sentences without an experiencer have yet to be tested in the literature, children will not be excluded on the basis of poor performance on this condition. While success on the unraised condition will therefore not be a prerequisite for inclusion in subsequent analyses (as is the case for the copula condition), those children who fail to comprehend the unraised sentences will be treated differently when it comes to analyzing responses to RNE sentences.

Turning lastly to the condition of primary theoretical interest, the RNE sentences (68) will hopefully help differentiate between the claims of Becker, UFH, and UPR. According to Becker, even young children have no difficulties comprehending StS raising, and therefore should respond correctly to the RNE sentences. Likewise, UFH predicts that children should have no difficulties with RNE sentences, as they do not involve exceptions to Freezing (i.e. unlike for ROE, no Smuggling is required for RNE). On UPR, however, children are expected to find RNE sentences ungrammatical, since just like with ROE sentences, RNE in the child grammar would involve movement across a strong phase boundary. Faced with an otherwise ungrammatical string, children might nonetheless seek (likely subconsciously) an alternative grammatical interpretation for RNE sentences. Once such interpretive strategy is the copula-analysis, where children treat the matrix StS raising verb as if it were the copula. Such an interpretation would take a verb denoting appearance (*seem* or *appear*) and replace it with a verb denoting predication ["reality"] (*is*). Since the manipulation of appearance and reality can now be confirmed on the basis of responses to copula and unraised sentences, there should be no doubt as to whether children make use of the copula-analysis in their interpretation of RNE sentences. If a child interprets *The dog seems to be purple* as *The dog is purple*, he will consistently respond incorrectly. Indeed, the copula-analysis should lead to statistically significant below-chance performance on the RNE condition. Such behavior is compatible with UPR and an interpretive strategy, but completely unexpected if children comprehend StS raising with no experiencer.

As for the details of how the actual test sentences are created, note that the adverb *really* is included in each test item, both to emphasize the difference between appearance and reality, and to draw the child's attention to the matrix verb, potentially discouraging him from ignoring the matrix verb entirely in favor of the embedded clause. The use of *really* is justified on the basis of our work with adult test subjects and pilot work with children. It demonstrated that its inclusion greatly increases the likelihood of correct responses to copula sentences. Furthermore, test sentences all include matrix verbs in the present tense. These test sentences are read while children look only at the second of the two test pictures, the first having been removed from view. The pairing of present tense with only the second picture serves to help isolate the moment at which the test sentence holds. This helps eliminate the possibility that justifications are being made on the basis of reference to different time points in the various scenarios across conditions.

Each condition is to be tested eight times, with four each of true and false items, for a total of 32 test sentences. The order of presentation for these items is pseudo-randomized such that children are not tested on the same condition more than twice in a row. Sentences are read twice before the child is asked to respond, and children's response justifications are written down before moving to the next test item. Sessions are also recorded onto audiocassettes. This is done in order to preserve the detailed child justifications, in case they are too long or complex to be fully transcribed by the experimenter in real time. By having children give justifications for their responses, we can attempt to gain greater insights into any errors and reasons for such. See Appendices (A6) for details.

Five native English-speaking adults, all of whom were naïve to linguistic theory and details of the experiment and its goals, were tested on this experiment, using methods identical to those to be used with the child participants. Each adult gave 100% correct responses on all conditions, establishing the validity of the experimental approach.

3.3.3 *Participants*

As to what age range to consider for inclusion among the child subjects in the current study, decisions had to be made as relates to the youngest and oldest ages allowable. Since all the previous StS raising studies discussed so far have found raising (ROE and RNE) to be acquired by the great majority of seven-year-olds, no children older than seven years of age were to be included. As for a minimum age of inclusion, despite Study 1 making clear that many pitfalls exist when studying StS raising in three-year-olds, since they fail to comprehend both *think* sentences (suggesting an issue with Theory of Mind, whether as relates to knowledge of belief states or their pictorial depiction) and the unraised form of StS raising verbs (suggesting either further ToM issues or a failure to recognize a particular lexical item as a StS raising verb), children as young as this were included in Becker's studies, so they will also be included here. The strong focus, however, will be on children aged four to five years-old, as this is the age range considered in Becker's second experiment. In order to maximize the validity of the obtained results, ten children in each one-year age range will be sought for inclusion in the study.

Other than age and requirement for native-English exposure, the only other inclusionary criterion was successful comprehension of the copula condition which is vital in order to draw any valid conclusions concerning children's comprehension of the RNE sentences. As such, comprehending the copula sentences is a prerequisite for further inclusion in the study. Any children failing to score at a statistically above-chance level on this condition (minimum 7 of 8 items correct) would be excluded and replaced with a child who did comprehend the copula sentences. The unraised condition is not being used to eliminate children in this study, as was the case in Studies 2, 3, and 4, since unlike in those previous studies, attempting to determine the percentage of children who do not comprehend unraised sentences is an actual goal of the current study. Those children who do fail to comprehend unraised sentences, however, will be treated separately for certain subsequent analyses, since it is hard to draw serious conclusions about their knowledge of RNE sentences.

Fortunately, every child tested comprehended the copula condition, and as such, no replacement subjects had to be found. These 50 children (27 boys, 23 girls) consisted of 10 children in each one-year interval between three and seven years of age (3.21-7.92 years, mean age 5.51). Participant details are shown in Table 3.3.1. All included children were normally developing native-English learners and came from families of varying socioeconomic status.

Age Group	#	Mean Age	Youngest	Oldest	Male	Female
3	10	3.65	3.21	3.97	3	7
4	10	4.37	4.03	4.90	6	4
5	10	5.47	5.07	5.99	5	5
6	10	6.54	6.12	6.93	7	3
7	10	7.55	7.13	7.92	6	4
Total	50	5.51	3.21	7.92	27	23

Table 3.3.1: Child participant details for Study 5.

3.3.4: *Results*

Accuracy results for all children across the four conditions are included below in Table 3.3.2.

Age Group	Copula	Control	Unraised	RNE
3	98.8%	100.0%	75.0%	2.5%
4	100.0%	100.0%	70.0%	36.3%
5	100.0%	100.0%	83.8%	33.8%
6	100.0%	98.8%	85.0%	67.5%
7	100.0%	97.5%	80.0%	71.3%
Total	99.8%	99.3%	78.8%	42.3%

Table 3.3.2: Accuracy for all conditions across all age groups for all children for Study 5.

Successful comprehension of the copula condition had been set as a precondition for inclusion in Study 5. Strikingly, all 50 children originally tested perform above chance on the copula condition (minimum 7 of 8 items correct), and as such, no replacements were necessary. Indeed, every four- to seven-year-old scores perfectly on this condition, with only one three-year-old missing a single copula sentence. Children have no issue distinguishing appearance from reality given these experimental materials on the basis of these data. Worth considerable attention is how different this result is compared to what was originally noted in pilot testing where the adverb *really* was not included in the copula test sentences. In cases where *really* was not included, children would often respond incorrectly, answering on the basis of *appearance* and not on the basis of *reality*, as is called for with copula sentences.

Just as with the copula condition, all 50 children score above chance on the syntactic subject control condition. Only three subject control test sentences (of 400) are answered incorrectly, and all age groups answer this condition at better than 97%

accuracy. These data serve to establish that indeed children have no trouble comprehending subject control.

On the unraised condition, children fare less well than is the case for the copula and subject control conditions. Still, all age groups score above 70% correct on the unraised sentences, and the older age (five to seven-year-old) groups all score at least 80% correct. Of the 50 children, 72% (36/50) comprehend the unraised items (answering a minimum of 7 of 8 unraised items correct). These 36 children, who comprehend the copula, subject control, and unraised sentences, are most relevant for addressing questions concerning the acquisition of StS raising (i.e. RNE) in this study. Worth noting, however, is that 28% (14/50) of the children tested do not comprehend the unraised sentences. This is important as it bears directly on Becker's claim that StS raising is acquired early. A claim of early StS raising acquisition appears much less tenable if over one in four children (in fact, 30% of the three- and four-year-olds, the ages Becker tested in her second experiment) cannot even comprehend the unraised variants of the raising sentences. This finding is in line with Study 1, where overall 19% of all children tested and 33% of children three to five years-old did not comprehend the unraised sentences.

Table 3.3.2 above also clearly reflects that the tested children have much greater difficulty with the raised sentences than with their unraised counterparts. Looking at the data from all children, while 78.8% of the unraised sentences are answered correctly, only 42.3% of RNE sentences are comprehended, an accuracy difference of 36.5%. While the great majority of all the children (72%) [36/50] are noted to comprehend the unraised sentences, the vast majority of all children younger than six years of age (86.7%) [25/30] fail to achieve above-chance comprehension on the raised condition.

Indeed, even if one only considers those children who scored above chance on the unraised condition, scores on the raised-condition are still extremely poor until around age 6 (Table 3.3.3). While six- and seven-year-olds, as groups, score quite well (both above 84% accuracy), four- and five-year-olds, again as groups, score right around chance level, while as a group, the three-year-olds score well below chance level.

Age Group	RNE	#
3	1.6%	8
4	45.8%	6
5	44.6%	7
6	85.7%	7
7	84.4%	8

Table 3.3.3: Accuracy scores by age group for RNE condition only considering those children who were also above chance on unraised condition in Study 5.

Extremely poor performance on the RNE condition is also reflected in an examination of how many children perform at a statistically significant above-chance level (Table 3.3.4). Only 36% (18/50) of all children tested respond at an above-chance level on RNE sentences, which is just marginally better when only considering those children who also responded at an above-chance level to the unraised condition. In this case only 50% (18/36) of the children can be taken to demonstrate comprehension of RNE sentences. Importantly, no child responds at an above-chance level on RNE who was not also above chance on the unraised condition. This suggests a dependency between first knowing the unraised sentences before being able to respond correctly to the RNE sentences. Of great relevance, failure to comprehend RNE sentences is greatest among the youngest children. Considering only the three- to four-year-olds (the relevant ages considered in Becker's second experiment), only 10% (2/20) comprehend RNE sentences at an above-chance level.[43]

Age Group	All Children	Know Unraised
3	0.0%	0.0%
4	20.0%	33.3%
5	30.0%	42.9%
6	60.0%	85.7%
7	70.0%	87.5%

Table 3.3.4: Percentage of children by age group who score above chance for RNE condition and for RNE once only including those children who are also above chance for unraised condition in Study 5.

While the above data demonstrate that the tested children, especially the younger ones, have great difficulty with correctly interpreting the RNE sentences, it is worth investigating exactly how they misanalyze these sentences. To this end, children in each age group are binned into clusters depending on whether they score at below-chance (0-1 correct), chance (2-6 correct), or above-chance (7-8 correct) level on the eight tested RNE sentences (Table 3.3.5).

Age Group	Chance Type	RNE
3	BC	10
	C	0
	AC	0
4[44]	BC	5
	C	3
	AC	2

[43] This (only) increases to 14.3% (2/14) if only considering those three- and four-year-olds who also score above chance on the unraised items.

[44] A mistake was made in Figure 2 in Hirsch, Orfitelli, and Wexler (2008) where one four-year-old is miscategorized (being marked as below chance when in fact he was above chance on RNE).

5	BC	7
	C	0
	AC	3
6	BC	3
	C	1
	AC	6
7	BC	2
	C	1
	AC	7

Table 3.3.5: Breakdown of chance performance for RNE condition across age groups in Study 5 (BC=below chance, C=chance, AC=above chance). The vast majority of children who are not AC respond at BC level, not C level. [Note: Includes children who are not above chance on unraised condition.]

It should be noted that the above table includes data from all 50 children originally tested, irrespective of whether or not they happen to comprehend the unraised sentences. While considerations of ultimate RNE acquisition are best made contingent on demonstrated mastery of unraised sentences, it proves beneficial here to consider all children when investigating overall response patterns due to the increase in usable data and the general irrelevancy of the unraised comprehension to issues of the particular patterns of RNE performance.[45]

In all, only 36% (18/50) of the children tested score above chance on the RNE condition. This clearly demonstrates that any claims about the early acquisition of RNE are misguided, since nowhere near 100% of the children score above chance on such sentences. Likewise, it is not the case that all children are blindly guessing when

45 It is interesting to note that even without knowledge of unraised sentences, these children appear to respond to the RNE sentences in line with the predictions of the copula-analysis (i.e. at below-chance level). Overall accuracy on RNE sentences for all children who are not above chance on unraised sentences is only 12.8%. Such a finding would of course be missed if one did not include those children who failed to comprehend the unraised sentences.

presented with RNE sentences, since on the binomial theorem, random guessing should result in only two children (1.76 exactly) scoring above chance, not 18. It is the case that some children do correctly interpret RNE sentences. These children tend to be older. Of the 18 children scoring above chance on the RNE condition, 72.2% (13/18) are six or seven years-old. Indeed, of the six- and seven-year-olds who comprehend the unraised sentences, 86.7% (13/15) also comprehend the RNE sentences.

Most of the children (64%) [32/50], however, do not comprehend the RNE sentences. Interestingly, it is not as if they are randomly guessing when presented with such sentences. Of those children not above chance on the RNE condition, only 15.6% (5/32) are at chance level, while the vast majority, 84.4% (27/32) respond at below-chance level. In fact, of those few children responding at chance level, 60% (3/5) score only one item away from being at below-chance level (i.e. they only get two of the eight RNE items correct). The other two children scoring at chance level get three and four RNE items correct, respectively. Taking the binomial theorem into consideration, these findings would be extremely surprising if children were randomly guessing on the RNE sentences. On a random binomial model, one expects to find only two children (1.76 exactly) scoring at below-chance level (instead of the actual 27 found here), while one expects 47 (46.5 exactly) children to be at chance level (instead of the actual five found here). The binomial theorem predicts 14 children scoring exactly four items correct if the 50 children are guessing randomly on RNE sentences, yet only one child is found to get precisely four RNE sentences correct. The data are strongly bimodal; children either respond at an above-chance level or at a below-chance level.

Clearly then there is very little random guessing going on here. Children basically fall into one of two camps: those that have acquired RNE (more than 70% of whom are at least six years-old) and those who consistently answer RNE sentences incorrectly (i.e. score below chance). It would, therefore, appear that those children who do not respond correctly to RNE sentences have a completely consistent interpretation for such sentences, an interpretation that just happens to lead to consistently incorrect (flipped) truth-value judgments.

3.3.5 *Discussion*

Of the 400 tested copula sentences, only one was answered incorrectly. This essentially perfect performance establishes that children have no general difficulties with either the task demands or the test materials. Importantly, the addition of *really* (along with the inclusion of pictures and the use of present tense) worked to distinguish the difference between appearance and reality for the children as its inclusion eliminated the overwhelming errors we found during pilot testing of the copula condition using Becker's experimental approach (see Section 3.3.1)

The subject control sentences are similarly well answered, with only three noted errors in 400 total test items. That is, children are basically perfect in their comprehension of subject control structures as tested here. No verb effect is found (*hate* and *like* are equally well understood). These data demonstrate that children are not ignoring the matrix subject control verbs (i.e. not using a copula-analysis), as this would result in chance performance. Furthermore, these data also rule out the possibility that children treat subject control verbs as (semantically vacuous) StS raising verbs, as this would result in below-chance performance. Children are most decidedly responding at an above-chance level to subject control sentences in Study 5. Indeed, children in this study perform much better than the equivalently-aged children in Becker's (2006) second experiment (Table 3.3.6). Since the experimental methodologies are generally quite similar between the two studies, this noted increase in performance on subject control sentences is likely due to the inclusion of the adverb *really*, the use of present tense, and the inclusion of pictures in Study 5.

Age Group	Becker	**Study 5**
3	65.9%	100.0%
4	88.4%	100.0%

Table 3.3.6: Accuracy for subject control condition across overlapping age groups for Becker's (2006) second experiment and Study 5.

In any case, of greater relevance, is that these results are much better than what Becker reports for her first experiment, on whose basis she originally claimed that subject control is delayed in acquisition. These results confirm that children do not treat syntactic control as StS raising. If children treated subject control sentences as StS raising sentences, they should perform at below-chance levels (i.e. consistently answer incorrectly). Yet, children are found here to provide consistently correct answers. What then to say about the data consistent with a control-as-raising analysis Becker found in her first experiment? It would appear this is nothing more than children willingly extending animacy to inanimates, that is, children taking the inanimate subjects to be sentient and thus licit controllers (see Section 3.3.1 for further discussion on this point). It is rather clear that indeed, subject control is an element of child grammar that is acquired at an early age.

Overall, unraised expletive-*it* sentences are generally well comprehended (78.8% accuracy across all children). Some of the younger children, however, have trouble with these unraised sentences, a finding that meshes with the previous studies documented in this dissertation investigating children's comprehension of unraised forms. The percentage of children who do not comprehend unraised sentences in Study 5 is fairly similar to what is noted for Study 1 (Table 3.3.3). Across the four to six year-old age groups, there is only a 10% difference in the rates of successful unraised comprehension (and only 20% for the seven-year-old group).[46] This is notable since the unraised sentences in Study 5, but not in Study 1, do not include an experiencer-phrase. No difference is found depending on which raising verb (*seem* or *appear*) is tested.

---

[46] Clearly there exists a rather large difference between the percentage of three-year-olds who comprehend the unraised sentences in Study 1 (40%) and Study 5 (80%), where many more three-year-olds in Study 5 are successful. An explanation for this is briefly presented below, but what is worth noting is that while the three-year-olds in Study 5 certainly know the unraised forms better than those in Study 1, this is not a challenge to any of the maturational theories under consideration, especially since, as noted below, these children are just as unsuccessful in their comprehension of StS raising.

That is, what concerns the maturational theories is genetically guided *development and acquisition* of syntactic knowledge (in this case StS raising). The question of whether a particular child has *learned* to correctly classify a particular verb (in this case as a StS raising verb) is certainly related to the former issue, but they are not the same issue (see Section 2.1.1 for further discussion). Lexical learning is ultimately largely dependent on non-genetically guided environmental factors (e.g. input frequencies, linguistic and situational context, etc). Indeed, Becker (2005, 2006) is crucially concerned with what factors eventually lead a child to categorize a verb as a StS raising verb (as opposed to a subject control verb).

The work presented here focuses more on questions related to the grammaticality of a particular set of syntactic operations (those involved in StS raising). The use of unraised raising sentences as an experimental control condition helps to establish a child's familiarity with a particular verb as a StS raising example so that successes and failures to comprehend raised sentences can more faithfully be interpreted (i.e. attempting to distinguish failure to have correctly categorized a verb vs. failure to have acquired particular syntactic operations).

There arises then the question of how children who lack StS raising nevertheless correctly respond to unraised sentences with StS raising verbs. How is it that children comprehend a verb as a StS raising verb (in its unraised form), but cannot interpret sentences with StS raising? This begs the question of whether verb categories themselves are innately specified, or whether children can only posit verb categories that conform to already posited syntactic operations.

Age Group	Study 1	Study 5
3	40.0%	80.0%
4	70.0%	60.0%
5	80.0%	70.0%
6	80.0%	70.0%
7	100.0%	80.0%

Table 3.3.7: Percentage of children in each age group above chance for unraised condition in Study 1 and Study 5.

While Becker (2006) reports three- and four-year-olds respond at 78.3% accuracy to RNE sentences in her second experiment, this result must now be tempered by the finding from Study 5 that 30% of similarly-aged children do not know the meaning of these raising verbs in their unraised form, raising serious questions about claims of children's early comprehension of raising with no experiencer. One should be very hesitant to argue that basically all young children have acquired RNE if a large minority of these children cannot correctly interpret the semantically equivalent unraised sentences.

Turning to the crucial RNE condition, children are found here to perform extremely poorly when presented with such StS raising sentences that do not include an experiencer-phrase. This is reflected both in accuracy scores and the percentage of children performing at statistically significant above-chance level. Again, no difference in performance is noted depending on which raising verb (*seem* or *appear*) is tested. Given that Becker's second experiment is nearly identical to Study 5, it is striking how the RNE results from each experiment are so different (Table 3.3.8). While Becker finds generally good, though certainly not perfect, comprehension of RNE sentences among her three- and four-year-olds (average accuracy of 71.2%), we find in Study 5 that children of the same age perform quite poorly (average accuracy of just 19.4%). The difference in children's RNE performance across the two experiments appears to be directly related to methodological problems in Becker's experiment that have been rectified in Study 5. Once children are required to demonstrate mastery of a copula condition (assuring they can make reality distinctions) and to demonstrate mastery of an unraised condition (assuring they can make appearance distinctions), it becomes quite clear that they do not comprehend RNE sentences, at least not until some time around the age of six or seven years-old.

Age Group	Becker	Study 5
3	64.0%	2.5%
4	78.3%	36.3%

Table 3.3.8: Accuracy for RNE condition across overlapping age groups for Becker's (2006) second experiment and Study 5. [Note: Results for Study 5 include children who do not comprehend the unraised condition since such children cannot also be excluded from Becker's data.]

While young children clearly have great difficulty correctly interpreting RNE sentences, it remains an open question as to how exactly they attempt to interpret such sentences. Insight into such interpretive processes, however, can be gleaned from children's particular response patterns to RNE sentences, as well as the justifications children use when responding to such sentences. As noted earlier, children are not randomly guessing when making truth-value judgments during RNE trials. Random guessing would be compatible with chance level performance, but fully 84.4% of children who do not comprehend the RNE sentences respond at below-chance, not chance, level to RNE sentences. Such performance suggests children are making use of a specific interpretive strategy when parsing RNE sentences, an interpretation that consistently maps to the wrong (opposite) meaning. One such strategy consistent with these data is the copula-analysis. If children systematically interpret the raising verb in a RNE sentence as the copula, presumably due to StS raising being otherwise ungrammatical for them (in line with UPR), this would result in consistently opposite truth-value judgments since the copula requires a judgment concerning reality in the scenarios while the raising verbs *seem* and *appear* require a judgment concerning appearance in the same scenarios. Since appearance and reality have specifically been put in complementary distribution in this particular experimental paradigm, confusing the two will result in consistently incorrect responses.

Further evidence that children are making use of the copula-analysis prior to RNE acquisition can be found in their justifications for their truth-value judgments. First, though, consider two typical correct responses (in this case "false") to the unraised test item (for two three-year-old children that are above chance on the unraised condition) in (69) and the children's justifications for their correct responses (with the children's ages noted in parentheses).[47] These two children are able to determine that the unraised sentence is false since Mr. Bear's perception of the scene does not match his statement. These children are clearly making their truth-value judgments as reflected in their justifications on the basis of how the scene appears to Mr. Bear (i.e. that the elephant looks to be in the shade), not based on the reality in the scene (i.e. that the elephant is actually in the sun).

[47] Also, note that none following justifications were "cherry picked" to make a point. Rather, these examples are just some of many that would all lead to the same conclusion.

(69) *It really seems that the elephant is in the sun.*

"It's dark with the sunglasses, so Mr. Bear thinks the elephant is really in the shade." (3;11)

"He's in the sun, but it looks like he's in the shade." (3;10)

Now consider the following semantically equivalent RNE test item (70), the exact same two children, and their response justifications where they incorrectly responded "true" when the correct response is "false". The justifications reflect only the reality in the scene (i.e. that the elephant is in the sun) and make no reference to how the scene appears to Mr. Bear (i.e. that the elephant looks to be in the shade). The children accept these sentences as being true precisely because they interpret the sentence as one affirming the reality of the predicate being applied to the subject.

(70) *The elephant really seems to be in the sun.*

"Cause he is in the sun. He's not really in the dark" (3;11)

"Because he's not in the shade. He's in the sun" (3;10)

Likewise, consider the following two RNE test items (71)-(72) and children's response justifications where they incorrectly responded "false" when the correct response is "true". Both examples demonstrate that the children are rejecting these sentences on the basis that the sentence does not match the reality in the scene (though both sentences of course are correct in that they match the appearance in the scene).

Neither of these children had any such difficulty responding on the basis of appearance with unraised test items.

(71) *The pig really appears to be in the water.*

"Because he's not floating in the water." (5;4)

(72) *The dog really seems to be purple.*

"No, he's not purple. In real life he's white" (3;10)

Many of the older, six- and seven-year-old children are able to respond correctly to RNE items, and crucially, their justifications do reference appearance, just as is the case for their justifications to unraised items. Consider the six-year-old who responds correctly to the test item in (73), where he answers "false" when indeed the item is false. This child directly acknowledges the reality of the scene (i.e. that the horse is actually small), but furthermore, he also recognizes that this is not the criterion on which to judge the truth of the sentence, and instead references the scene's appearance in rejecting it.

(73) *The horse really seems to be small.*

"He is small, but the magnifying glass makes him look big" (6;4)

This pattern of response justifications to RNE sentences among younger children, whereby judgments are based on reality and not appearance, along with the consistent incorrect (i.e. below chance) responses is completely expected if children are subject to the copula-analysis when interpreting RNE sentences. The incorrect justifications (confusing reality for appearance) disappear at the very time that children begin to respond correctly to the RNE sentences (between six and seven years of age).

If the young children in this study cannot interpret RNE sentences in an adult manner (the hypothesis being for grammatical reasons), is it really the case that they all basically adopt an interpretation consistent with the copula-analysis, as opposed to an interpretation along the lines of the subject control, and hence, reflexive *expect*-analysis, as appears to be the case for many children in Study 4? Just as with the RNE sentences in Study 4, the RNE sentences in Study 5 all involve animate subjects (a prerequisite for a subject control), so subject animacy is not an issue. It is the case, however, that in Study 5 (unlike Study 4), the belief state of the grammatical subject in each test sentence is not at issue. Rather, what matters is only the belief state of Mr. Bear, who is never mentioned in any test sentences. It would therefore seem unlikely that children would adopt an interpretation on which truth-judgments are based on the thoughts of the subject in the RNE sentences, when that subject's thoughts are not detailed or focused upon in the experimental setting.

That said, it is the case that adopting the *expect*-analysis for RNE sentences would nevertheless likely lead to below-chance performance, consistent with the noted pattern of truth-value judgments. Consider the test sentence *The pig seems to be in the water.* Given the scene, this sentence is true. The copula-analysis gives (roughly) *The pig is in the water*, which is false (the pig is indeed on land, it only appears, via his reflection, that he is in the water). The *expect*-analysis would give a parse meaning *The pig thinks he is in the water.* While the scenario never focuses on where the pig believes he is standing (on land or in the water), it is certainly conceivable that the child might conclude that the pig knows where he is standing, even if that fact is unclear to Mr. Bear. In this case then, the sentence is false (the pig thinks he is standing on land). This reflexive interpretation, therefore, leads to the same incorrect truth-value judgment as given by the copula-analysis. Crucially, however, the particular response justifications make it clear that those children who respond incorrectly to RNE sentences are doing so in accord with the copula-analysis, whereby they cue on reality and not appearance, and not with the *expect*-analysis, since no justifications cue on the subject's thoughts or belief state. It is here that gathering response justifications proves helpful over simply recording the child's judgments.

Ultimately, the copula-analysis is the more general interpretive strategy (at least with RNE sentences). It can, unlike the *expect*-analysis, be extended to StS raising sentences with inanimate subjects, thereby also capturing the results from Becker's first experiment. In addition to being consistent with the response pattern and response justifications in Study 5, it is also consistent with the results from Study 2, and is needed to explain why so many more children in Study 4 perform better on RNE sentences than ROE sentences. The copula-analysis, however, cannot be extended to all StS raising sentences, and alternative analyses seem to arise where the copula-analysis cannot apply. For example, with a medial experiencer-phrase, the copula-analysis provides an ungrammatical parse for ROE sentences (e.g. **John is to Mary dancing*), and children's interpretations instead appear consistent with a RtO interpretation (e.g. the *imagine*-analysis; *John imagines Mary to be dancing*). When the thoughts of the grammatical subjects are put into focus, as is the case in Study 4, some children (fully 50% of four-and five-year-olds tested) also do not use the copula-analysis, but instead make use of the subject-control *expect*-analysis (borne out both by their response pattern and response justifications).

What appears to be the case, in summary, is that children are not tied to a particular interpretive strategy when presented with the raising verbs in question (*seem* and *appear*) in a StS raising frame, but rather, are flexible (to a limit) in extending meaning to sentences containing these verbs. These raising verbs, at least in raising sentences, do not have a meaning consistent with a pre-defined interpretation, otherwise the varied results from Studies 1-5 cannot be captured. Children are obviously quite intelligent, and appear to seek parses that allow for meaningful interpretation.[48,49] That different linguistic and environmental contexts result in different parses, therefore, is not all that surprising. It would seem then that consistent lexical interpretation can only follow once the relevant grammatical operations have been acquired (in this case, StS raising).

---

[48] Here *meaningful* encompasses being both correct (obviously) and non-random (less obviously). The ideal interpretation is of course one that maps correctly between linguistic inputs and the real world. An interpretation that leads to consistent chance performance (akin to random guessing) is really rather useless. The *expect*-, *imagine*-, and copula-analyses all map to incorrect responses in Studies 1-5. Importantly, however, they allow for non-random (though ultimately incorrect) responses. To the extent that they allow non-random behavior, however, they are quite useful (i.e. "actionable" as discussed in Section 2.3.7). Ultimately, the child will acquire (via maturational development) the relevant syntax to allow for correct adult interpretations, and the continued failure of the compensatory heuristics will lead to adopting the adult parse.

[49] Intelligent language learning is not the exclusive domain of children acquiring their native language. The positing of intelligent hypotheses also arises in second language (L2) acquisition. It might very well be possible to learn more about how and why children posit the compensatory heuristics they do by investigating the use of such strategies in L2 contexts. An interesting experiment would be an examination of how speakers of an L1 that does not allow StS raising over a full DP experiencer (e.g. Italian) would learn a language that does allow such ROE sentences (e.g. English). There is no reason not to expect errors that are consistent with the RtO *expect*- or *imagine*-analyses, but of course, this has (by hypothesis) nothing to do with maturation, and everything to do with lexical learning and interpretive strategies. For example, unlike with children (see Study 3), there is no strong prediction that Italian-English L2 learners should have positively correlated scores for StS raising and verbal passives.

Before turning to the implications of these data to the question of which grammatical theory best captures them, a few comments are warranted on the inclusion and role of the adverb *really* in this present study. Some critics might attempt to pin the difference in RNE performance between Becker's second experiment and Study 5 solely on the inclusion of *really* in the latter study. This, however, would be a rather gross simplification. First, note that the presence of *really* in both the copula and subject control conditions results in absolutely no difficulties for the children (both conditions answered at above 99% correct). Indeed, it is precisely the addition of *really* in the copula condition that led to children successfully responding to this condition, compared to when it was omitted in pilot testing. Furthermore, the inclusion of *really* in the RNE (and unraised) sentences proves utterly unproblematic for the adult control subjects, basically all of the seven-year-olds, and the majority of the six-year-olds. In order to claim that *really* is the cause of children's poor performance with RNE sentences, one would have to defend (and explain) a complex interaction account where *really* interacts only with StS raising to cause a problem in just the three- to five-year-olds. Finally, if *really* were the issue (and it is not), one would have to predict no correlation between RNE comprehension and verbal (subject experiencer) passive comprehension (Study 6). The use of *really* in Study 5 is not a problem; rather, it is a crucial part of the solution to issues with Becker's original experiments.

These data, namely poor performance on RNE sentences in Studies 4 and 5, also speak to the possibility of a parameter missetting explanation (see Section 3.1.2) for children's noted difficulties with ROE in Studies 1, 2, and 4. Such an account would tie children's failure to comprehend ROE sentences to an initial (default) parameter setting whereby only RNE is licit, such that the relevant evidence necessary to license ROE had not yet been discovered. Since RNE is also not licit in early child grammar, however, such an explanation is no longer tenable. Maturational theories of language acquisition seem more plausible given the current experimental data set.

An examination of StS raising comprehension across age groups in the various studies of this dissertation shows a relatively common developmental trajectory, as measured by the percentage of children per age group who respond at a statistically significant above-chance level, once controlling for knowledge of the raising verbs as reflected in unraised sentences (Table 3.3.8).[50] In Studies 1, 2, 4, and 5, which all include children aged four to seven, it is only with the oldest age group, the seven-year-olds, that the majority of the children in the age group comprehend the StS raising sentences in every study. For these same four studies, four- and five-year-olds have great difficulty with the raising sentences (in neither age group do at least half of the children respond above chance in any of the studies).

Age Group	**Study 1**	Study 2	**Study 4**	**Study 5**
4	16.7%	20.0%	0.0%	33.3%
5	12.5%	26.7%	0.0%	42.9%
6	25.0%	40.0%	30.0%	85.7%
7	66.7%	80.0%	70.0%	87.5%

Table 3.3.8: Percentage of children per age group who perform above chance on StS raising sentences in Study 1, Study 2, Study 4, and Study 5 for children who are also

[50] Within age group percentage of children above chance is a better measure for comparison across various studies than accuracy, since the latter is more heavily affected by different interpretive strategies, where errors in one study might arise as chance performance, while in another as below-chance performance. For example, assuming two strategies that both result in incorrect responses for all children, but where one strategy leads to chance guessing (assuming two choices: 50%) and the other leads to below-chance responses (0%), the accuracies are quite different, but a measure of the percentage of children above chance for both cases would match exactly (0%). Both measures are of course equally susceptible to noise introduced by interpretive strategies that result in above-chance performance even in the face of linguistic immaturity. Such strategies are devilish for language acquisition researchers.

above chance on unraised condition. [Note: Studies 1, 2, and 4 all involve ROE, Study 5 involves RNE. Only age groups shared across the four studies are included.]

Interestingly, while six-year-olds as a group in Studies 1, 2, and 4 perform poorly, never realizing a majority of children who comprehend the raising sentences, the six-year-olds in Study 5 do quite well (85.7% above chance; twice the percentage as found in any of the other three studies). One is left to wonder what is special about the six-year-olds or the StS raising sentences in Study 5. First, it should be noted that the data in Table 3.3.8 reflects ROE for Studies 1, 2, and 4, and RNE sentences for Study 5. When the RNE data from Study 4 is considered, six-year-olds are again noted to do quite well (70% above chance). Is this evidence that while RNE is certainly delayed, perhaps it is not delayed as long as ROE? In the case of the RNE data from Study 4, it is at least possible that many of the six-year-olds scoring well actually lack StS raising syntax, but are making use of the copula-analysis, which just happens to map to the correct answers in that study. It is the case that the six-year-olds in Study 5 are on average slightly older than the six-year-olds in Studies 1, 2, and 4. It could also be that Study 5 randomly includes more linguistically mature six-year-olds than normally expected and reflected in the other studies. This possibility is addressed in Study 6 by investigating whether the six-year-olds who score well on RNE sentences also happen to score well on subject experiencer passives, which are usually acquired around seven years of age.

As Studies 1, 2, and 4 demonstrate, ROE is delayed in early linguistic development. Furthermore, ROE acquisition appears to be tied to verbal passive acquisition, on the basis of the data in Study 3. These results are equally compatible with

UPR and UFH, two grammatical maturation theories. In order to decide between these two grammatical accounts, data concerning the acquisition of RNE sentences was needed. UPR predicts that RNE sentences will also be ungrammatical for premature children, while UFH makes no predictions about such sentences. Study 5 basically replicates and extends the finding from Study 4 that RNE is delayed in acquisition. As such, UPR receives tremendous empirical support, while UFH is basically refuted. UPR also predicts that RNE and verbal passives should follow similar acquisition trajectories, an expectation that is taken up in Study 6.

Some comments are warranted on the appeal of the StS raising experiment put forth in Study 5. It of course owes its creation to Becker. Yet it includes many improvements; notably it adds the copula condition and the unraised condition that together provide an independent test of children's ability to discern a reality/appearance dichotomy. Furthermore, the inclusion of the unraised sentences also permits a lexical control for StS raising,: it allows the experimenter to distinguish possible cases where StS raising sentences are not comprehended due to grammatical problems with raising, versus cases in which such sentences are not comprehended due to lexical learning problems, namely that the child does not know the particular raising verb used. This experiment also results in better unraised performance compared to that in Study 1. In Study 1, 45% of three- and four-year-olds did not comprehend the unraised sentences, compared to only 30% having similar difficulties in Study 5.[51] The experimental task in Study 5 also requires fewer ToM considerations, since the child need not track the belief state of every interacting character (as was the case in Study 4). Fully 50% of the three-year-olds in

[51] This could possibly be due to the structural simplicity of the unraised sentences in Study 5 (no experiencer-phrase) compared to those in Study 1 (with experiencer-phrase).

Study 1 had to be dropped due to ToM issues (failure on the *think* trials), while no three-year-olds in Study 5 had issues with the copula or subject control sentences, where successful comprehension of either requires some ToM considerations.

Most importantly, however, Study 5 involves a task and linguistic materials that are not at all dependent on having an experiencer-phrase in the StS raising test sentences. This avoids the crosslinguistic issues raised in Section 3.1.2 and avoids the processing issues detailed in 2.3.1. Removing the experiencer completely allows a much more straightforward investigation of StS raising, as opposed to merely raising over an experiencer, which comes with sone baggage. Finally, the scenarios in Study 5 are much faster to run than those in Study 4, since the latter require manipulating three puppets (which requires two experimenters) and a host of assorted inventory items. The real world benefit for researchers is that more children can successfully be run on the much shorter experiment in Study 5. Future studies investigating StS raising comprehension, both in typically-developing children and in other populations, should seriously consider adopting this methodological approach.

3.3.6 *Conclusion*

Becker (2005, 2006) claims that RNE is early in acquisition (in place by age three). Her relevant (second) experiment is predicated on the idea that if tested, children would also respond correctly to a copula sentence expressing the same predicate as tested with the RNE sentences, essentially showing that children are sensitive to a reality-appearance dichotomy. Our pilot work attempting to replicate Becker's finding of early RNE comprehension, however, finds that if such copula sentences are explicitly tested, children get them wrong, which throws into doubt any claims about successful RNE comprehension using such an experimental methodology.

Making several small, but crucial changes to Becker's second experiment (adding the adverb *really* to all test sentences, consistently using only present tense morphology in the test sentences, and making use of two pictures to more concretely distinguish reality from appearance) results in children successfully comprehending the relevant copula sentences. The inclusion of unraised sentences also helps establish an awareness of the reality-appearance dichotomy, and provides a direct lexical control for the StS raising verbs. Such additions to the experimental paradigm ultimately reveal that children are delayed in their comprehension of RNE sentences. Indeed this delay matches well that found in the previous studies (Studies 1, 2, and 4), where StS raising is acquired by most children somewhere between six and seven years of age.

Young children's failure to comprehend RNE sentences rules out the possibility that earlier noted difficulties with ROE sentences could be due to a form of parameter missetting, such that children believed they spoke a language that only ruled out raising over an experiencer, since it is clear now that all StS raising (with or with no experiencer) is equally delayed. This result also has the effect of ruling out UFH. While UFH correctly predicts delays for ROE sentences, it incorrectly expects children to have no difficulties with RNE sentences. UPR, however, predicts exactly this pattern of results, namely that all StS raising should be ungrammatical for young children.

UPR also makes the strong prediction that StS raising and verbal passives should be similarly delayed (same developmental trajectories and within-subject acquisition correspondence). This prediction was confirmed for ROE and verbal passives in Study 3.

In Study 6, the prediction will be tested for RNE and verbal passives. Such a test will hopefully also shed some light on why six-year-olds performed so well on the RNE sentences in Study 5.

3.4 Study 6: Relationship Between Raising with No Experiencer and Verbal Passives

3.4.1 *Motivations*

On the basis of the results obtained in Studies 4 and 5, it has become clear that in addition to StS raising over an experiencer being delayed in language acquisition (see Chapter 2 for details), StS raising with no experiencer is similarly delayed. This result is exactly as predicted by UPR, since both ROE and RNE involve (in the adult grammar) movement across a defective phase boundary, an operation illicit on UPR. While UFH is able to capture the child data concerning comprehension of ROE, where an exception to the Freezing Principle is needed (in the adult grammar), it cannot explain children's poor performance with RNE, since the latter does not involve Smuggling.

UPR furthermore predicts a correspondence between the acquisition of verbal passives and RNE. Since both structures are ruled ungrammatical on UPR for the same reason, namely that they both involve (*A*-)movement across a (for the child) strong phase boundary, they should both be delayed to the same extent. This entails similar developmental curves for both structures, and more to the point, within-subject correspondence for the acquisition of each structure. To be precise, children should fall into only two groups: those that have acquired both structures (i.e. having matured; no longer subject to UPR) and those who find both structures ungrammatical (i.e. having not yet matured; subject to UPR).[52] UFH does not predict delays for RNE, and thus there is no expectation of a correspondence between the acquisition of RNE and verbal passives on UFH. A finding that RNE and verbal passives track one another in language development would be another strong piece of evidence in favor of grammatical maturation theories in general, and UPR specifically.

Finally, in Studies 1, 2, and 4, children were generally found to acquire StS raising (ROE) around seven years of age. It is thus initially somewhat surprising that in Study 5, StS raising (RNE) is acquired at a slightly earlier age (six years-old). While the age difference is relatively minor (just one year), it begs a few questions. First, is StS raising (ROE and RNE) generally acquired around age six or age seven, or is there a large enough standard deviation with respect to age of acquisition to make such a specific question irrelevant? Second, is it possible that RNE is actually acquired earlier than ROE? While Studies 4 and 5 offer compelling evidence that RNE is delayed (contra Becker and UFH), could it nevertheless be the case that ROE is delayed further than RNE? Third, is it possible that Study 5 simply includes a number of developmentally precocious six-year-olds? It might very well be the case that in general, StS raising (both ROE and RNE) is usually acquired around age seven, but that by chance, Study 5 ended up with a disproportionate number of already linguistically mature (with respect to UPR) six-year-olds.[53] It is here that the within-subject testing of RNE and verbal passives can

[52] This claim is actually a bit too strong given that morphological and lexical learning also play important roles in syntactic interpretation.
[53] Statistically speaking, if most children are expected to acquire StS raising at seven years of age, then finding so many six-year-olds to have acquired it (i.e. ROE) in Study 5 should probably be a bit surprising. How surprising is of course a function of the population standard deviation for age of StS raising acquisition, the details of which are unknown. Indeed, the data of this dissertation basically provide the first glimpse of what such a standard deviation might look like. Ultimately, more raising acquisition data is

help tremendously in answering these various questions. As noted in Section 1.2.2, across many studies, children are generally found to acquire subject experiencer verbal passives around seven years-old. If RNE and verbal passive acquisition are indeed developmentally linked, the six-year-olds who performed well on RNE in Study 5 should also perform well on subject experiencer passives (here in Study 6), even though six-year-olds typically do not do well on such sentences.

3.4.2 *Experimental Design*

This study consists of two experimental tasks, each of which is administered to the same children, so as to offer within-subject data on when both RNE and verbal passives are acquired. The RNE task is exactly that from Study 5 (see Section 3.3.2). This is basically a TVJ task in which children must respond to four conditions, consisting of copula, subject control, unraised expletive-*it*, and RNE sentences. The verbal passive experiment is exactly that from Study 3 (see Section 2.4.2). It involves a two-choice picture selection task and a two-factor design, crossing verb type (actional vs. subject experiencer) and voice type (active vs. passive). The crucial comparison will pit acquisition of RNE sentences with subject experiencer verbal passives. As was the case in Study 3, for most analyses, only those children who comprehend the unraised sentences from the raising task and the active sentences from the passive task will be included.

---

needed before one can make informed determinations about what constitutes surprisingly early acquisition of StS raising. Of course, it is worth noting that some children as young as three years old comprehend StS raising sentences in the various studies, so no one is claiming that a child must be seven years old before acquiring this grammatical structure.

3.4.3 *Participants*

The same 50 children who took part in Study 5 were also administered a test of verbal passive (and active sentence) comprehension. Each child was required to complete both experiments within 30 days, with the raising experiment run first and the passive experiment second. The average time span between the tests being administered is only five days. In fact, for 30 of the children (60%) the span was zero days, meaning that both experiments were completed on the same day.[54] In less than a week on average (and a month at most), it is extremely unlikely that the relevant grammatical maturation could have taken place such that a child fails on the first raising task but succeeds on the later passive task. Participant details are shown in Table 3.4.1.

Age Group	#	Mean Age	Avg Span	Min Span	Max Span
3	10	3.64	11	0	29
4	10	4.37	2	0	10
5	10	5.47	10	0	28
6	10	6.53	0	0	1
7	10	7.55	1	0	7
Total	50	5.51	5	0	29

Table 3.4.1: Child participant details for Study 6.

3.4.4 *Results*

The results of the passive experiment appear below in Table 3.4.2. This table includes data for all 50 children. Every child performed at a statistically significant above-chance level on both active conditions (i.e. regardless of verb type). The basic findings replicate previous studies to cross voice and verb type, finding an interaction

[54] The six- and seven-year-olds have less of a time gap between the two experiments being administered. This is basically just a factor of the older children having more focused attention, allowing both experiments to be completed more quickly, often on the same day.

between these two factors, where both active conditions are comprehended well, but in which subject experiencer passives are much more poorly comprehended than actional passives. This is understandable on the hypothesis that children lack the grammatical operation(s) necessary to derive verbal passives, but can nonetheless make use of a compensatory linguistic parse (the adjectival strategy) to capture the (general) meaning of actional passives, but not subject experiencer passives (see 1.2.3 for detailed discussion).

Age Group	Act. Actives	SE Actives	Act. Passives	SE Passives
3	97.5%	100.0%	81.3%	41.9%
4	96.3%	98.8%	86.3%	59.4%
5	98.8%	100.0%	96.3%	65.6%
6	97.5%	96.3%	95.0%	88.1%
7	96.3%	100.0%	96.3%	84.4%
Total	97.3%	99.0%	91.0%	67.9%

Table 3.4.2: Accuracy for all conditions across all age groups for all children for the passive experiment portion of Study 6. [Act. = Actional, SE = Subject Experiencer]

For the purposes of the remaining analyses, only data from those children who comprehend the unraised sentences in the raising test will be included. Successful comprehension of unraised items is a natural prerequisite for faithfully interpreting results on raised sentences, and allows valid conclusions to be drawn about any developmental relationship between RNE and verbal passives. This inclusionary requirement rules out 14 children, leaving usable data for 36 children. While Study 3 includes data from 55 children, once the eight- and nine-year-olds are ignored, this leaves data from 35 children, which basically matches Study 6.

Now one can ask how the data from the latest passive experiment compare to the previous passive experiment in Study 3. The relevant condition involves the subject experiencer (SE) passives, since these passives are not subject to the adjectival strategy, and therefore provide a better look at children's true ability to comprehend verbal passives. Accuracy across age groups for subject experiencer passives in both Studies 3 and 6 are presented in Table 3.4.3. Both studies find generally weak performance in children aged three to five years, though the four- and five-year-olds in Study 6 are doing somewhat better than the equivalently-aged children in Study 3 (see Footnote 56). Both studies find seven-year-olds to have very little difficulty with these verbal passives. Six-year-olds in Study 6 do quite well on subject experiencer passives, while those in Study 3 have many more difficulties, a point to which we return below.

Age Group	Study 3[55]	#	Study 6[56]	#
3	43.8%	4	40.6%	8
4	45.8%	6	70.8%	6
5	57.0%	8	70.5%	7
6	48.4%	8	92.0%	7
7	80.6%	9	88.3%	8

Table: 3.4.3: Accuracy for SE passives across age groups once children who did not score above chance on unraised condition have been removed for Study 3 and Study 6.

[55] Note, these are not the same values as found in Table 2.4.1, since that table did not exclude children who failed to comprehend the unraised sentences from the raising task.

[56] The SE passive accuracy scores for the four- and five-year-olds in Study 6 are poor (where 75% accuracy is above chance for 6 of 8 items), but they are relatively better than what was noted in Study 3. It is initially surprising that both of these age groups perform around 70% correct on SE passives in Study 6 given that such children have elsewhere been noted not to generally comprehend SE passives until around the age of six or seven (Section 1.2.5). One might wonder if collectively children in Study 6 are scoring slightly better than those in Study 3, or whether there are simply more children scoring above chance in Study 6, thereby pulling up the group average.

The Study 6 passive findings appear mainly to be a reflection of a slight difference between studies in terms of both the number of children in each of these two age groups to comprehend the unraised sentences, and the number of children to actually comprehend the SE passives. While 42.9% (3/7) of the five-year-olds in Study 6 are above chance on SE passives, this number falls to 25% (2/8) in Study 3. If only examining those children who did not score above chance on SE passives, one finds very similar and near-chance (50%) SE passive accuracies: Study 3 – 44.4% (6 children), Study 6 – 53.1% (4 children). Similar numbers hold for the four-year-olds: Study 3 – 36.3% (5 children), Study 6 – 59.4% (4 children).

A similar analysis can be made on the basis of the percentage of children in each age range who comprehend subject experiencer passives (Table 3.4.4). Here again, one sees that a majority of children within an age group comprehending the SE passives is only achieved with the seven-year-olds in Study 3 and with the six-year-olds in Study 6.

Age Group	**Study 3**	#	**Study 6**	#
3	25.0%	4	0.0%	8
4	16.7%	6	20.0%	6
5	25.0%	8	42.9%	7
6	25.0%	8	85.7%	7
7	66.7%	9	87.5%	8

Table 3.4.4: Percentage of children above chance for SE passives across age groups once children who did not score above chance on unraised condition have been removed for Study 3 and Study 6.

Turning to the question of whether or not RNE and verbal passives track one another in language development, Figure 3.4.1 plots the developmental curves for RNE and verbal (SE) passives as a function of age group accuracy over a four year age span. These two developmental curves are remarkably similar in shape, with the offset in the three-year-olds plausibly being accounted for by a difference in compensatory heuristics, where these youngest children are left to guess randomly for SE passives (50%) and employ the below-chance copula-analysis for RNE sentences (0%); see Footnote 56 for comment concerning the offsets among the four- and five-year-olds.

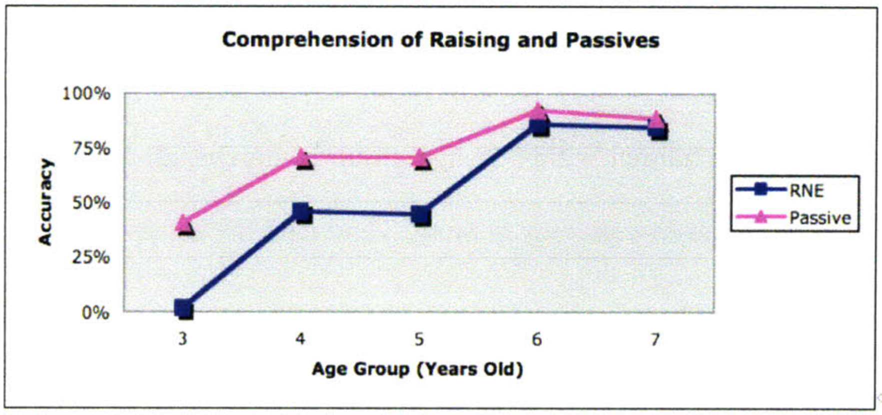


Figure 3.4.1: Similar developmental curves for comprehension of RNE and verbal (SE) passives as measured by age group accuracy for those children scoring above chance on unraised condition in Study 6.

Further developmental curves can be generated by plotting the percentage of children in each age group who perform at an above-chance level for RNE and verbal (SE) passives across the four age groups (Figure 3.4.2). Strikingly, these two developmental curves overlap perfectly.

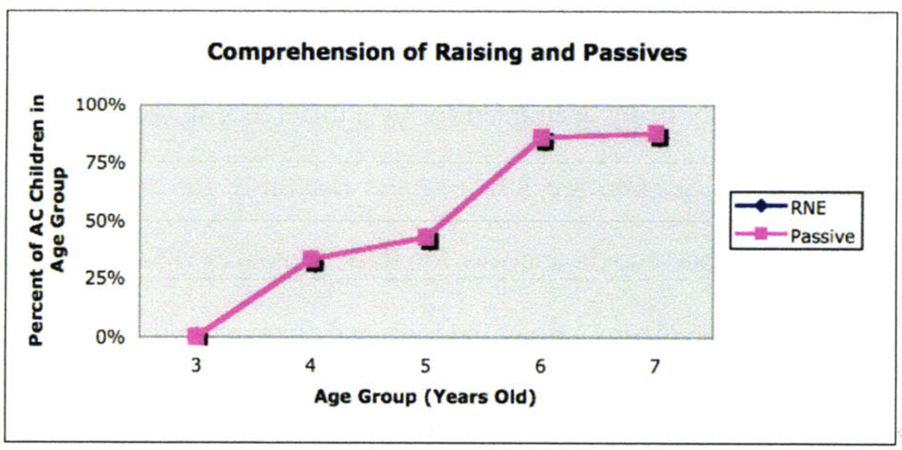


Figure 3.4.2: Similar developmental curves for comprehension of RNE and verbal (SE) passives as measured by children in each age group scoring at above-chance level on unraised condition in Study 6.

The data for these children can also be summarized in a scatter plot, plotting each child's accuracy on RNE against his accuracy on verbal (SE) passives (Figure 3.4.3). Cutoffs for statistically significant above-chance performance are included as orange lines in the scatter plot: 75% in the case of passives and 87.5% in the case of RNE.[57] This figure shows that, in general, older children (six- and seven-year-olds) tend to cluster in the upper-right quadrant (above chance on both structures), while children younger than six years-old tend to populate the lower-left quadrant (not above chance for either structure). Crucially, and in strict accordance with UPR, not a single child falls in either the upper-left or lower-right quadrants. Children either comprehend both structures or they comprehend neither.

[57] Here, statistically significant above chance performance for RNE is at least 87.5% correct (7 of 8 items correct). For passives, statistically significant above chance performance is defined as at least 75% correct on both truncated and full subject experiencer passives (6 of 8 items correct on both subconditions), which is relatively conservative.

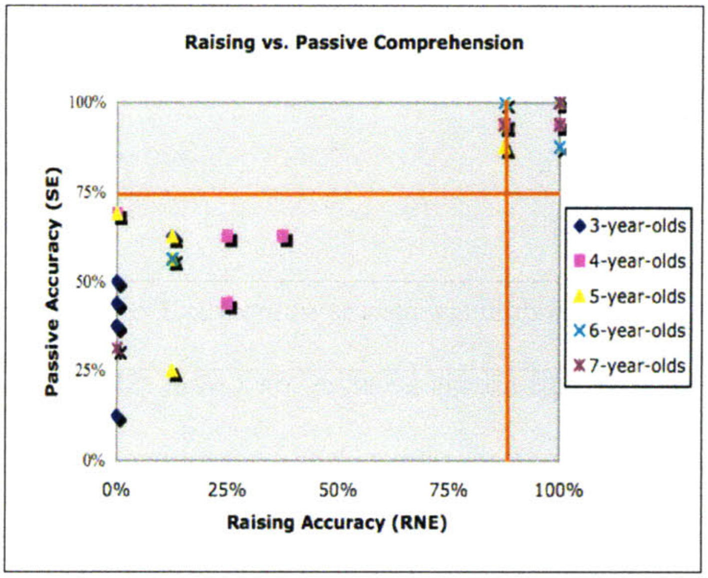


Figure 3.4.3: Scatter plot for comprehension scores from SE passives and RNE for 36 children from Study 6. Notably, two groups of children are observed: a group in the lower left quadrant who do not know either structure, and a group in the upper right quadrant who know both structures. Orange lines indicate a statistically significant above-chance level of comprehension for each structure. [Note: Many data points overlap.]

Finally, one can statistically examine the strength of the relationship that holds between these two structures by calculating a Pearson (r) correlation. Statistically significant, and very high correlations for RNE and passives obtain when either children's exact scores ($r(34) = 0.917$, $p < 0.0001$) or above-chance performance ($r(34) = 1.000$, $p < 0.0001$) is examined.[58]

---

[58] As noted in Study 3 (Chapter 2, Footnote 79), using a Pearson correlation with binary (above- or not above-chance) data is an unintended use of the statistical test. A better analysis, one that basically shows the same effect, involves simply computing a correspondence between the two structures, noting whether for each child above or not above chance performance on the two structures matches or mismatches. Dividing the total number of cases of match (here 36) by the total number of children (here 36) provides the

3.4.5 *Discussion*

The results from the passive experiment basically repeat those from previous studies that have crossed verb type and voice: both actional and subject experiencer actives are acquired early, while actional passives are delayed, with subject experiencer passives acquired last (Section 1.2.2). The passive data gathered here, though, do show two slight deviations from previously reported findings. First, the youngest children (three-year-olds) in Study 6 perform markedly better on actional passives than in many previous studies (e.g. Study 6: 81.3% vs. Study 3: 65.6%, including data from all three-year-olds). As suggested by Hirsch, Modyanova, and Wexler (2006), and discussed in Section 1.2.5, actional passives, but not subject experiencer passives, are subject to environmental influences, and it might simply be that the children in Study 6 have been raised in environments conducive to earlier actional passive acquisition. Note of course, that many of these very children are not without linguistic delay, however, as they have great difficulty with RNE sentences (Study 5) and subject experiencer passives.

That said, the second difference noted in this study compared to previous investigations of passive acquisition is that many of the younger children here do rather well on the subject experiencer passives. The three-year-olds in Study 6 are not remarkable in this respect, responding basically just as poorly on subject experiencer passives (40.6% correct) as in Study 3 (43.6% correct). The next three age groups, however, all perform noticeably better in Study 6 compared to Study 3. Only considering those children in each study who correctly comprehend the unraised sentences of the

---

correspondence (here 1.0). There is perfect correspondence between comprehension of RNE and verbal passives in Study 6.

corresponding raising study, Study 3 shows a weighted average of 51.1% correct for subject experiencer passives among four- to six-year-olds, while the same age groups in Study 6 have a weighted average of 70.7% correct for the same passives. This difference is perhaps most striking in the six-year-olds, who still perform overall generally poorly in Study 3 on subject experiencer passives (48.4% correct), while children of the same age do quite well in Study 6 on these passives (92.0% correct) using exactly the same test materials. While such a difference is striking and worth further discussion, which follows below, the central question of Study 6 is not merely one of attempting to replicate previous acquisition findings concerning verbal passives, but to investigate the relationship between the pattern of acquisition of verbal passives and RNE sentences. While the maturational theories predict a similar time course of acquisition for these structures across various studies (i.e. groups of children should generally acquire these structures at roughly similar ages), ultimately, their strongest prediction is that both structures should be acquired at the same time (i.e. that within-subject acquisition of both structures should be positively correlated).

Indeed, the maturational theories receive tremendous support from Study 6, as RNE and subject experiencer passive acquisition appears inextricably linked. The acquisition curves for both structures are remarkably similar. In the case of examining accuracy scores on each structure, the curves have the same general shape, with an offset of about 25% among the younger children. This offset is accounted for by the posited compensatory heuristics being applied to deal with each of these structures by children who have yet to acquire the necessary syntax for adult comprehension. In the case of the subject experiencer passives, those children not comprehending these sentences are left guessing, which leads to chance (i.e. 50% given the two-picture choice) performance. In the youngest group, basically all children are blindly guessing (40.6% correct). Among the four and five year-olds, there is a mix of children who comprehend the structures (i.e. 100% correct; relevant syntax acquired) and those who do not (i.e. 50% correct; chance guessing), leading to group performance in the 70% range. Performance accuracy on the RNE sentences is different for those children who have failed to acquire StS raising. Here random guessing is not seen, instead, young children who have not acquired StS raising appear to make use of the copula-analysis, which leads to consistently below-chance responses on RNE sentences. Thus, the youngest group here is basically around 0% accuracy. The four and five year-olds are again a mix of children who comprehend the test sentences (i.e. 100% correct) and those who do not (i.e. 0% correct), leading to group performance in the 50% range. Crucially, however, while an offset, one which is understood given the different strategies for dealing with otherwise ungrammatical sentences, is present when comparing RNE and SE passive accuracies across age groups, the shapes of the developmental curves themselves match nearly perfectly. This is impressively reflected when considering the acquisition curves for these two structures for above-chance comprehension, a dependent variable that is (relatively) immune to compensatory strategies. In this case, one finds exactly overlapping developmental curves for RNE and SE passive sentences.

This strong group level relationship between the age of acquisition of RNE and SE passives is rendered more meaningful when examining the within-child comprehension of both structures. As seen in the scatter plot in Figure 3.4.3, the group level relationships established in the developmental curves in Figures 3.4.1 and 3.4.2 are arising due to a precise correspondence between the comprehension of RNE and SE passives at the individual child level. Children either comprehend both structures or they comprehend neither. Development of each therefore appears strongly linked.

Such a strong developmental linkage is unexpected on many of the hypotheses under consideration. Certainly this relationship cannot be explained by Becker, who offers no account for verbal passive delay, nor for RNE for that matter. The correspondence in acquisition is quite unexpected on any explanation that attempts to attribute delays in either raising or verbal passive comprehension to having to learn each structure on a verb-by-verb basis (e.g. Gordon and Chafetz, 1990). It seems rather doubtful that a non-linguistically-specific working memory developmental theory could be spun that ties processing load of RNE and SE passives (to the exclusion of subject control and many other structures of apparently similar linguistic complexity) together. Of great relevance, UFH does not predict these results. While verbal (SE) passives are predicted to be delayed on UFH, RNE sentences are not. UFH offers no explanation for the strong group level and individual child level developmental linkage of these two grammatical structures. These data, along with those in Study 5 demonstrating delay for RNE sentences, strongly argue against UFH as an explanation for delayed acquisition of SE passives and ROE sentences.

These findings, however, are in accordance with the predictions of one already discussed grammatical theory: UPR. Both RNE and SE passive sentences involve the same syntactic operation (i.e. movement across a weak phase) that is posited to be delayed for children subject to UPR. Once children mature linguistically and UPR no longer constrains their grammar, both RNE and SE passives become licit structures, and children quickly move away from the compensatory strategies that previously applied when attempting to interpret these structures.[59] This maturation appears to take place some time around age six.

Finally, the data gathered in Study 6 help in understanding the unexpected early acquisition of RNE in Study 5. The prior StS raising studies (Studies 1,2, and 4) all find comprehension to improve markedly around age seven, in general accordance with when verbal passives are acquired. Study 5 finds StS raising comprehension (in this case for RNE) to improve at age six, a full year earlier than the other raising studies. On the basis of Study 5 alone, it remains unclear whether to take these data as indicative that RNE is acquired (a year) earlier than ROE and verbal passives, and thus not in accordance with the predictions of UPR, or whether Study 5 simply includes an abundance of linguistically precocious (i.e. mature) six-year-olds. If RNE acquisition indeed generally precedes ROE acquisition by a year, then one would expect to see in Study 6 a divergence in the acquisition of RNE and SE passives in six-year-olds. Alternatively, if ROE, RNE, and verbal passive acquisition are all linked due to a common grammatical delay (e.g. UPR), then one expects to see the six-year-olds in Study 6 to similarly respond well to verbal passives, just as they did to RNE sentences in Study 5 in divergence with findings from prior verbal passive experiments. That the six-year-olds in Study 6 do better on SE passives that what is found in most previous SE passive acquisition studies

[59] Presumably this transition occurs since none of the posited compensatory heuristics leads to consistently above-chance comprehension in all situations. What role, if any, negative evidence plays in speeding this transition once the relevant linguistic maturation has taken place remains unclear. It is hard to imagine, however, that a child could keep applying indefinitely an interpretive strategy (e.g. *expect-* or copula-analysis) that leads to below-chance comprehension without *some* corrective feedback, nor does it seem feasible that a child would continue to maintain a strategy (e.g. as in the case of SE passives) that always results in random guessing once an "actionable" interpretation becomes available.

does suggest that Studies 5 and 6 simply included a number of linguistically mature children (see Footnote 53).

3.4.6 *Conclusion*

Study 6 replicates prior experimental work demonstrating an interaction of voice and verb type, with comprehension of SE passives being delayed to a greater extent than actional passives. Developmental curves for the acquisition of RNE and SE passives are generated and found to be quite similar. This holds both of accuracy scores and number of children within each one-year age interval. An examination of within-child comprehension of both grammatical structures finds a perfect correspondence: the acquisition of one structure is a perfect predictor for the acquisition of the other. The two structures appear to be delayed for a common reason.

These data cannot be accounted for under UFH. While UFH correctly predicts delayed acquisition of verbal passives and ROE sentences, it does not capture the finding of delayed acquisition of RNE sentences (Studies 4 and 5), nor can it explain why RNE and verbal passive acquisition mirror one another. UPR, however, correctly explains the linked acquisition or verbal passive, ROE and RNE sentences. Given the strong evidence in support of a maturational account of verbal passive acquisition (Chapter 1) and the strong linkage between the acquisition of verbal passives and StS raising presented in Study 6, it stands to reason that StS raising is also subject to maturation. The relevant linguistic target of maturation appears to be the allowance of weak phases and subsequent agreement and movement operations that follow.

In order to judge the ultimate validity of UPR as an explanation for the findings under discussion, further considerations are warranted. First, child production data of StS raising sentences ought to be considered. Second, there exist concerns that warrant comment about the grammatical viability of UPR given current knowledge of syntactic theory. Finally, one must consider whether UPR makes further correct predictions concerning the acquisition patterns of other, yet to be discussed syntactic structures. These three concerns are all investigated in the next chapter.

3.5 *Summary and Subsequent Directions of Inquiry*

The studies presented here in Chapter 3 have attempted to address how children interpret StS raising structures when no experiencer-phrase is included. These RNE sentences are delayed just like ROE sentences from Chapter 2, once the relevant masking interpretive strategies are controlled. RNE appears to be delayed to the same extent, and thus likely for the same reason, as verbal passives.

3.5.1 *Study 4*

We replicate the earlier studies of the previous chapter and demonstrate that unraised structures with StS raising verbs are generally acquired quite early, while their ROE counterparts are acquired quite late (around age six or seven). As was the case in Study 1, children again appear to interpret ROE sentences with animate subjects in accordance with a *think*-for-*seem* interpretation (i.e. the raising-to-object *imagine*- and *expect*-analyses). Most of the tested young children (younger than age six) also have great difficulty with RNE sentences, also analyzing them with a *think*-for-*seem* interpretation (i.e. the subject control *expect*-analysis).

Difficulties with RNE sentences is evidence against UFH, as well as against the idea that early child grammar might be similar to those languages that allow RNE but not ROE (e.g. Spanish, Icelandic). Rather, these data are compatible with UPR, which predicts that all StS raising should be delayed. Yet, more children correctly respond to RNE than ROE sentences, which is unexpected on UPR. This result, however, is compatible with UPR if only some children adopt the *expect*-analysis, while others adopt the copula-analysis (of Study 2), where the latter children respond correctly, but for non-adult reasons. What is needed therefore is an experiment that can rule out the application of the copula-analysis, since then UPR predicts problems with RNE to match those noted for ROE in previous studies. Becker's (2005, 2006) claims that RNE might be early in acquisition also need to be addressed.

3.5.2 *Study 5*

This study begins with a critical review of Becker's (2005, 2006) conclusions that StS raising is not subject to delay in linguistic development. First, evidence that subject control is acquired early, certainly earlier than verbal passives and ROE (and RNE from Study 4), undermines Becker's claim that children interpret subject control structures as involving StS raising. Second, possible methodological issues are raised for her studies suggesting RNE is early in development.

Those issues are addressed in a new experiment. A new copula-condition is added to ensure that children differentiate appearance from reality. A new unraised condition is also added to ensure that any difficulties with the RNE condition are not simply due to children's unfamiliarity with the StS raising verb (as opposed to StS raising syntax). Study 5 demonstrates that once the copula-analysis is controlled for, Becker's results are not replicated, and that RNE is just as delayed as ROE from previous studies. Problems with RNE sentences constitute strong support for UPR, and strong evidence against UFH: all StS raising is delayed, not just ROE.

3.5.3 *Study 6*

While Study 5 offers compelling evidence that RNE sentences are delayed just as has been observed for ROE sentences, in support of UPR, the strong prediction of UPR (to which UFH does not speak) is that RNE acquisition should correspond with verbal passive acquisition. This relationship should hold not just of age groups, but within individual children. On UPR, all StS raising structures and verbal passives are delayed due to a common deficit (the inability to represent weak phases, thus allowing movement from the complement of *v*).

Testing the same children on verbal passives and RNE reveals that both structures are acquired at the same general age (six to seven years old), both structures have matching developmental curves, and crucially, that there is a perfect correspondence between the acquisition of the two structures for individual children: those children who comprehend one comprehend both, or else they know neither.

3.5.4 *Challenges and New Directions*

The six studies detailed in Chapters 2 and 3 offer much data suggesting that both ROE and RNE are delayed in typical language acquisition. Whatever accounts for this delay also appears to hold for verbal passives, since the same children who are delayed on verbal passives also have difficulties with StS raising (ROE and RNE), and vice versa. As such, maturation grammatical theories (e.g. UPR) that predict both structures to be delayed gain great support.

The extent to which UPR is a valid account of these data still remains an issue. One must also consider UPR's predictions as relates to both production data and other grammatical structures (e.g. unaccusatives, control, etc). Only if UPR continues to make accurate predictions about which structures are delayed/early in development can it remain a viable explanation. Also, UPR, like any grammatical account of language acquisition, relies on certain assumptions concerning adult grammar. These assumptions must hold in order for it to work (e.g. existence of strong/weak phases, lack of Smuggling in ROE, etc). If these assumptions prove false, UPR becomes untenable. It is to such concerns that we turn next in Chapter 4.

Chapter 4: Further Data, Further Considerations[1]

The previous two chapters focused in particular detail on children's comprehension of subject-to-subject raising sentences, and how such performance related to children's comprehension of verbal passives. The data gathered suggest that StS raising is not only delayed in acquisition (until sometime around six years of age), but that it is delayed for reasons similar to those governing the late acquisition of verbal passives. Of the various grammatical accounts considered, the Universal Phase Requirement appears to best capture the various data. It alone correctly predicts that verbal passives and StS raising structures with (ROE) and without (RNE) experiencer-phrases will be similarly delayed.

In this final chapter, we examine further evidence that bears on issues surrounding children's acquisition of StS raising. In the first sub-section (Section 4.1), the explanatory adequacy of UPR relative to the various data is considered in further detail. In part, child-directed and child-produced natural production data are examined to add to the comprehension data gathered from prior experimental investigations. Furthermore, predictions made by UPR as they apply to other grammatical structures are considered. Next, the particular syntactic framework that UPR necessitates is discussed. In the last sub-section (Section 4.2), implications of the conclusions reached are highlighted as they relate to bigger questions in language acquisition, linguistics, cognitive sciences, and biology. Finally, future directions of inquiry are briefly laid out.

[1] Study 7 originally appears in Hirsch and Wexler (2007a), but further data and analyses are considered here for the first time. Ken Wexler contributed heavily to the interpretations and conclusions that I draw here, but is obviously in no way at fault for any errors I may have reached.

4.1 Further Data and Grammatical Considerations

The comprehension data gathered in Studies 1-6 demonstrate a delayed acquisition of StS raising, irrespective of the presence of an experiencer-phrase, which appears to be fundamentally linked with delayed acquisition of verbal passives as has been previously described in the language acquisition literature. Such data are compatible with UPR, a grammatical account that posits children lack defective verbal phases for maturational reasons. This account differs from the Universal Freezing Hypothesis, on which premature children's grammar does not allow the syntactic operation of Smuggling due to a strong version of the Freezing Principle. According to UFH, while children should demonstrate a comprehension delay for StS raising with an experiencer-phrase (as found in Studies 1-2), they should have no problem with StS raising without an experiencer-phrase. Contra UFH, however, and in accord with UPR, children are likewise delayed on StS raising without an experiencer-phrase (as found in Studies 4-5). The various data examined in these studies also serve to rule out other grammatical accounts (i.e. ACDH, EARH, CAH) that had attempted to capture children's delayed comprehension of verbal passives. Only UPR correctly predicts a similar and correlated delay for StS raising (with and without experiencer-phrase) and verbal passives.

Of critical importance, however, is the ultimate explanatory adequacy of UPR for the noted experimental results. While UPR appears to best capture the hitherto discussed findings, it must be asked to what extent is UPR the very best possible theory for the data. That is, could another, yet unarticulated, grammatical account be the actual correct explanation? The ultimate soundness of UPR rests on its ability to make further correct predictions in regard to the acquisition timeline of other grammatical structures that require defective verbal phase in adult grammar, and in the end, the validity of the syntactic machinery on which it relies and rests. To the extent that one can find acquisition evidence that goes against the predictions of UPR, or against the syntactic operations or representations required by UPR, an alternative account will have to be sought. To help decide the validity of UPR, further data, beyond those already gathered, are required. Some such data are examined below, while others will have to await future investigation.

4.1.1 *Study 7: Adult and Child Productions Containing* Seem

Past acquisition studies of subject-to-subject raising, including Studies 1-6 discussed in Chapters 2 and 3, have failed to investigate the degree to which children hear StS raising structures and StS raising verbs, as well as the extent to which children themselves produce such structures and verbs. Since part of the central interest in this work is whether young children comprehend StS raising (and if so, how it is interpreted), it is obviously desirable to know whether children are even exposed to such structures and whether they produce them. In order to determine the degree to which children hear and produce StS raising structures, we examined the child-directed and child-produced speech for all 1051 English-speaking children on the CHILDES corpus containing the raising verb *seem* (MacWhinney, 2000).[2] Detailed analyses were limited to utterances containing the verb *seem* after initial searches made clear that *seem* was by far the most

[2] At the time the searches were conducted (5/2004), this constituted every English-speaking child available in the CHILDES database.

frequent StS raising verb in the corpus, and would thus serve as the sole verb in the comprehension experiments detailed in the previous two chapters.[3]

There are several specific research questions that such an examination can help address. First, production data can help speak to the grammaticality of StS raising for young children. While the already-gathered comprehension evidence strongly suggests that StS raising is ungrammatical for young children, it could be significantly harder to conclude that it is indeed ungrammatical if one were to find copious use of such structures in early child utterances. Furthermore, by paying attention to children's use of experiencer-phrases, it is possible that competing grammatical accounts could be differentiated. For example, many productions of StS raising sentences without experiencer-phrases, but few with experiencer-phrases, might be taken as evidence for UFH. One the other hand, a general lack of any StS raising sentences would better fit the predictions of UPR. In addition, careful analysis of any child-produced StS raising sentences might offer evidence in favor of a particular compensatory heuristic. For example, examining both the animacy and theta-roles of the embedded subject in such utterances could help distinguish the *expect*-analysis from the copula-analysis, or even the soundness of the *imagine*-analysis. Finally, by taking a close look at child-directed speech, one might be able to deduce explanations for any comprehension or production delays, as well as possible heuristic strategies, and ultimately, how children go about differentiating raising and control verbs.

To such ends, the following factors were coded and analyzed for all child-directed and child-produced utterances containing the verb *seem* in the CHILDES corpus: the

[3] While these analyses focus exclusively on *seem*, it should be noted that other raising verbs do appear in the input to these children, including dozens of examples of *appear*, *tend* (to), *used* (to), and *happen* (to). The verb *seem*, however, is the most frequently used and heard by children, by a large degree.

child's exact age, the utterance itself (including URL in case later context was needed), whether it was a raised or unraised form, the animacy of the embedded subject, the theta-role of the embedded subject, the presence or absence of a *to*-phrase experiencer, and whether or not the embedded clause was adjectival or verbal.

From 552 child-directed utterances containing *seem*, 448 analyzable, non-repetitive utterances were extracted for further investigation. For the sample considered (a little under a million utterances), this implies that a child hears a unique sentence containing *seem* every 1700 utterances. While this might seem scarce, an average American child hears about 7000 utterances per day (Cameron-Faulkner, Lieven & Tomasello, 2003). Extrapolating, this means that by a child's third birthday (the youngest age considered in the comprehension studies in Chapters 2 and 3), he will have heard 4500 sentences containing *seem*. It is worth noting that the number of input examples containing *seem* in this data set exceeds the number of input utterances containing verbs that most researchers (and parents!) would judge young children to know, including *crawl*, *feed*, *hug*, *lift*, *pass*, and *rub*. Thus, it is unlikely that any difficulties children might have in comprehending raising structures involving *seem* could be attributed solely to the rarity of the verb in the input, as the verb simply is not that rare. Also, StS comprehension difficulties could not be attributed directly to children failing to hear the verb used in raised constructions. The vast majority of parental use (87%) is in the raised, non-expletive form (Figure 4.1.1).

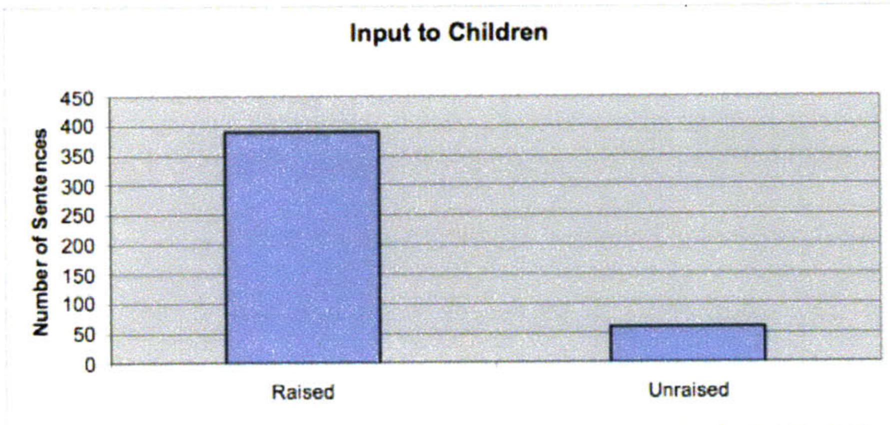


Figure 4.1.1: Bar graph showing the number of analyzable raised and unraised *seem* sentences directed at children in the CHILDES database. The vast majority of such input sentences are of the raised type.

Furthermore, it is not the case that adults in some way modulate their use of unraised and raised structures, such that the raised sentences are only used with older children, as an explanation for why only older children (in general) comprehend StS raising sentences. The type of structure used by the adults (raised or unraised) is not a function of the child's age; raised forms are used at all ages, even with children who have not yet reached their first birthday (Figure 4.1.2).

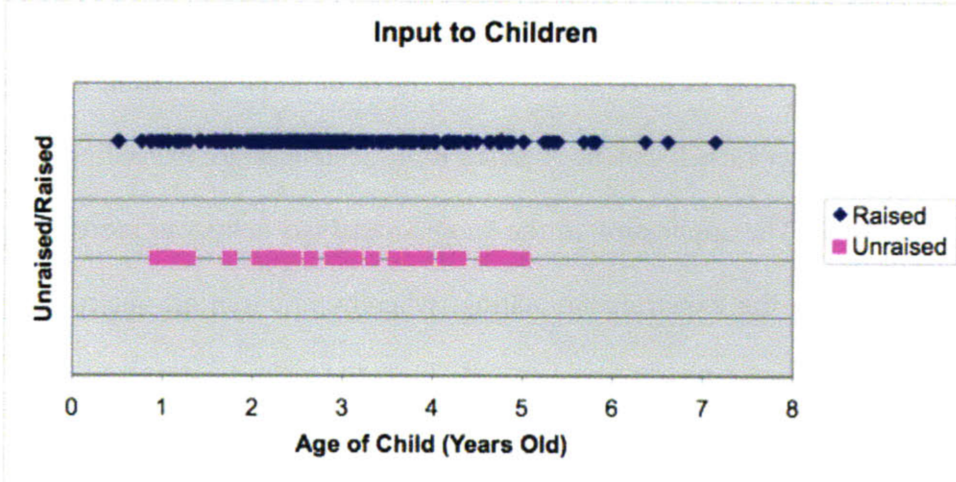


Figure 4.1.2: Plot of all analyzable raised and unraised *seem* sentences directed at children in the CHILDES database by child's age.

Overall, only 28 (6.3%) of the child-directed *seem* sentences contain an experiencer-phrase. Of these, the vast majority occur with unraised sentences. While a hefty 23.7% of adult unraised sentences contain an experiencer, only 3.6% of adult raised sentences do. Of the 28 adult sentences containing an experiencer, 22 occur in sentence-final position (e.g. "well it seems a bit unsafe to me"), six in sentence-medial position (e.g. "it seems to me to be rather continental"), and zero in sentence-initial position. Turning only to the adult raised sentences, 225 (57.8%) involve an animate subject.[4]

The above data help address ideas about the compensatory heuristics children were hypothesized to produce when interpreting otherwise ungrammatical raised sentences. The *imagine*-analysis posits that children interpret StS raising sentences with medial experiencers (ROE) as raising-to-object sentences. Yet only two such sentences are found in the corpus. While the ROE comprehension data are quite consistent with the

[4] A good many of the *seem* sentences involve either a demonstrative or pronominal subject. In such cases, context derived from the previous and subsequent utterances was used to determine subject animacy.

imagine-analysis, even across multiple experimental paradigms, it would seem that such an analysis is unlikely to have arisen by means of children tabulating the appearance of medial experiencer raised sentences, since ROE sentences are exceedingly rare in child-directed speech. The application of the *imagine*-analysis is instead more likely occurring spontaneously in the experimental setting, whereby children are interpreting otherwise new-to-them strings (*seem* sentences with a medial experiencer-phrase) by analogy to RtO sentences, which *are* common enough in the input (there are hundreds of child-directed RtO sentences in the corpus). The adult raising sentences without experiencer-phrases (RNE) also help address the possible origin of the copula- and *expect*-analyses. Since subject control requires an animate (sentient) subject, its use as a heuristic with raised *seem* sentences rests on many examples of raised *seem* appearing with animate subjects. Indeed, over half of such sentences do have animate subjects. The fact that there is a good mix of animate and inanimate subjects with these sentences might give rise to apparently different compensatory strategies in Study 4. Those children who have heard many adult raised *seem* sentences with animate subjects might choose the subject *expect*-analysis, while those who have heard a good number of such sentences with inanimate subjects might be more likely to choose the copula-analysis.

The greater number of child-directed raised sentences compared to unraised sentences (by a nearly 6.6:1 ratio), might help account for why certain of the youngest children in various comprehension studies had difficulties with unraised sentences. As previously noted, children must correctly learn to distinguish raising and control verbs. In addition to the fact that subjects do not have to be animate, the presence of an unraised variant is a piece of unambiguous evidence that a particular verb is a raising verb and not a control verb. The rarity of unraised forms in the input to children means some time will pass before a particular child hears enough utterances to correctly classify *seem* as a raising verb. Until the verb is so classified, one cannot ask whether or not the child can comprehend a raised sentence involving that verb. It is thus perhaps not surprising that the youngest children tested had greater difficulty with unraised sentence forms. Importantly, however, once *seem* is learned (as measured by comprehension of unraised sentences in the various studies), young children still have great difficulty comprehending raised sentences, which is further evidence against a simplified learning theory where lexical knowledge directly predicts syntactic knowledge.

While the child-directed speech is quite interesting, especially as relates to understanding how various heuristic strategies might form, it is ultimately the child-produced utterances that are of particular significance, as these bear most strongly on the issue of child StS raising grammaticality. First, it must be asked if children produce both unraised and, more importantly, raised StS sentences. Second, to the extent that raised sentences are produced, what form do they take and at what ages do they appear? UPR predicts that while unraised sentences should not be delayed for maturational reasons (whether or not they are delayed for learning reasons is left open), all StS raised sentences, whether with an experiencer or not, should be delayed. The mere presence of child-produced raised sentences, however, cannot alone be taken as evidence against UPR, as both child age and the possibility of alternative syntactic analyses for such sentences must be addressed. It is with respect to the latter issue that coding for subject animacy, theta-roles, complement type, and presence of experiencer-phrase come into play.

While these input analyses make clear that children do indeed hear raising sentences with the verb *seem*, analysis of children's use of such structures shows that children rarely produce them. In all, only 67 child-produced utterances containing *seem* appear in the corpus, of which only 33 constitute non-repetitive, analyzable examples. Of these 33 utterances, 21 (64%) are raised sentences and 12 (36%) are unraised (Figure 4.1.3). The percentage of raised uses compared to unraised uses is thus marginally less in the children than the adults.

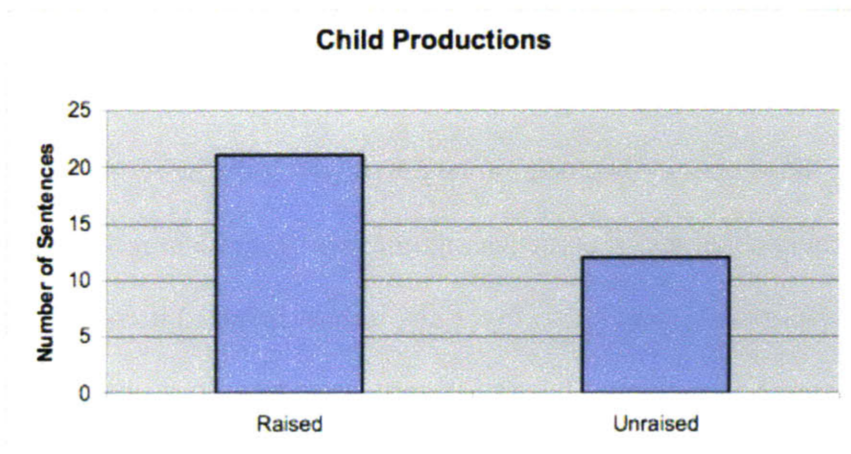


Figure 4.1.3: Bar graph showing the number of analyzable raised and unraised *seem* sentences produced by children in the CHILDES database. The majority of such child-produced sentences are of the raised type.

What is much more interesting than the above finding is that there is an almost 14:1 ratio of adult-use to child-use of sentences containing *seem*, while the overall ratio of adult-produced to child-produced unique utterances in the corpus is less than 2:1. This input-output discrepancy with the verb *seem* is even more striking given that for verbs with similar input frequencies, such severe discrepancies do not exist (e.g. 1.9:1 for *carry*, 2.1:1 for *climb*, 3.4:1 for *crawl*, 2.3:1 for *dance*, 0.8:1 for *feed*, 0.7:1 for *hug*, 4.0:1 for *lift*, 2.1:1 for *pass*, and 4.1:1 for *rub*). Furthermore, there appears to be an age effect, such that of the analyzable child-produced examples, only 11 (33%) come from children younger than five years old, while the majority (67%) are produced by children five years old and older. The extreme paucity of child-produced *seem* sentences, especially in light of their relative abundance in child-directed speech, does at least suggest that sentences with StS raising are somehow marked in child grammar.[5]

Of the 21 raised sentences produced by the children, only two involve an experiencer-phrase, and only one has the experiencer-phrase appearing sentence-medial (the other is sentence-final).[6] The absolute rarity of child-produced ROE sentences (namely, a single example) fits extremely well with the grammaticality predictions of both UFH and UPR. The one ROE sentence is produced by a five-year-old. While this is at the low end of the age range at which children appear to acquire StS raising, it is worth noting that no children younger than this produce even a single such sentence.[7] UPR predicts RNE sentences to be ungrammatical for immature children, while UFH hypothesizes them to be licit in early child grammar. The extremely small number of RNE sentences (only 19) would seem to argue against UFH, in particular relative to the

[5] The fact that raised *seem* sentences are generally absent from child productions is obviously in line with the predictions of various grammatical accounts. What accounts for the lack of unraised *seem* sentences, however, is not as clear.

[6] Only a single child-produced unraised experiencer-phrase sentence was noted. It is rather interesting that children generally performed as well as they did on the unraised sentences with experiencer-phrases in Studies 1, 2, and 4, given how few they produce in their natural speech.

[7] While both UPR and UFH do predict ROE sentences to be ungrammatical, it was shown in both Study 1 and Study 4 that children tend to interpret such sentences in accord with the *imagine*-analysis, namely, as raising-to-object sentences. The small number of child-produced ROE sentences strongly suggests that the *imagine*-analysis, while a powerful comprehension heuristic, remains something children only seek out when presented with otherwise uninterpretable sentences. That is, children do not come to the experimental paradigms with the idea that *seem* is a RtO verb, else one would expect to see it used more often in natural child production.

total number of child-produced utterances and the number of RNE sentences in the input to these children. What, though, to make of these 19 child-produced RNE sentences and what might this mean for UPR?

First, it is worth noting that while UPR does predict such sentences to be ungrammatical, it could be that any productions are merely reflections of a compensatory heuristic, such that what appear to be RNE sentences actually do not involve any StS raising. Of the 19 such sentences, six (32%) involve animate subjects. Of those six, four are agents and two are themes; none are experiencers. *If* these 19 RNE sentences are the product of a compensatory heuristic, the above breakdowns mean that the utterances are only compatible with the copula-analysis, and not the *expect*-analysis. The latter analysis involves subject control, which is neither compatible with inanimate subjects (68% of the child-produced RNE sentences have inanimate subjects) nor with non-experiencer subjects (100% of the child-produced RNE sentences have non-experiencer subjects). Thus, the extremely few child-produced raised sentences could simply be a product of a compensatory heuristic, and therefore not involve any StS raising, such that they are not counter-examples to UPR. An alternative account for these few child-produced raising sentences comes about when looking at the type of complement clause present relative to the child's age.

A closer examination of children's raised sentences reveals a striking asymmetry between those produced by children younger than 5.5 years of age and those 5.5 years and older. Of the twelve examples produced by the younger children, eleven (92%) involve small-clause adjectival complements (e.g. *That seems* <u>*fun*</u>; 3;0), while eight of the nine older children's utterances (89%) have verbal complements (e.g. *They seem* <u>*to be*</u>

following the same direction; 7;0). It is thus possible that younger children's grammar allows *seem* with adjectival small-clause complements, but not with fully verbal complements, which might be due to different syntactic analyses depending upon with which type of complement *seem* appears.

Why should there be a difference in the age at which *seem* with VP complements (verbal-*seem*) develops versus *seem* with small-clause (adjectival) complements (adjectival-*seem*)? Note that there are considerations suggesting that adjectival-*seem* does not involve the type of StS raising relevant to UPR. Adjectival-*seem* (1) tends not to allow an experiencer *to*-phrase between itself and its complement, while there is no such problem with verbal-*seem* (2).[8]

(1) *John seems to Bill sad.

(2) John seems to Bill to be sad.

The experiencer, though, is licit with adjectival-*seem* if it is fronted (3), but a fronted experiencer is also allowed in copular constructions (4).

(3) To Bill, John seems sad.

(4) To Bill, John is sad.

[8] The restriction on the experiencer-phrase does not seem limited to adjectival complements, but rather, any small-clause complement, whether an AP (i), PP (ii), or NP (iii).

(i) *John seems to Bill sad.
(ii) *John seems to Bill in the backyard.
(iii) *John seems to Bill a detective.

All of the above are fine (at least certainly much better) with a clearly verbal complement clause.

(iv) John seems to Bill to be sad.
(v) John seems to Bill to be in the backyard.
(vi) John seems to Bill to be a detective.

When taking a verbal complement with a stage-level predicate like *available*, verbal-*seem* allows both existential (5a) and generic readings (5b).

(5) Firemen seem to be available.

a. There exist some firemen x such that x seem to be available (existential reading)

b. For all firemen x, x seem to be available (generic reading)

With adjectival-*seem*, however, the existential reading (6a) is unavailable, while only the generic reading (6b) is present:

(6) Firemen seem available.

a. *There exist some firemen x such that x seem available (existential reading)

b. For all firemen x, x seem available (generic reading)

In Diesing's (1992) framework, the existential reading is derived via lowering the subject from [Spec, IP] back into [Spec, VP]. This, of course, is only possible when the subject initially raised from [Spec, VP] to [Spec, IP]. For the generic reading, on the other hand, the subject is directly generated in [Spec, IP], forming a subject-predicate relationship, with a control analysis. The lexical DP in [Spec, IP] controls a PRO subject in [Spec, VP], which is assigned a theta-role by the verb. This suggests an analysis like (7).

(7) John seems sad.

[$_{IP}$ DP John [$_{I'}$ INFL seems [$_{VP}$ PRO sad]]]

Consequently, the subject cannot lower and only the generic interpretation is licensed. Thus, adjectival-*seem* is syntactically very similar to the copula. Crucially, there is no defective *v* involved, in fact no raising of the argument at all, thus no violation on UPR. As such, young children's early productions with adjectival-*seem* are no challenge to UPR, which actually predicts their early use, alongside the much later development of verbal-*seem*. While this analysis could certainly use refinement (and further relevant production data would be nice), it suggests an interesting set of future investigations centered around the projection status of the embedded clause with raising verbs.

These production data paint a picture where children are exposed to many *seem* sentences, but nevertheless produce very few on their own. The little that are produced tend to fit with already-posited compensatory heuristics, or are explained once further syntactic analyses are undertaken. These production data are compatible with predictions (including UPR) that StS raising is ungrammatical in young child grammar.

4.1.2 *Unaccusatives and Control as Raising*

A good part of the desire for wanting to investigate children's knowledge of StS raising stems from the predictions about its time course of acquisition made by various grammatical acquisition theories originally posited to account for verbal passive acquisition. Many of those theories were equally compatible with the (verbal passive) acquisition evidence available. It is by turning to a new grammatical structure for which those acquisition theories also make predictions, in this case StS raising, that their validity can be tested. Likewise, the ultimate validity of UPR is also dependent on it making correct predictions for grammatical structures other than verbal passives and StS raising sentences. While UPR ultimately makes interesting predictions for many grammatical structures, some of which it is hoped will be explored in future studies, this section will limit itself to discussion of two particular syntactic phenomena: unaccusatives and control as raising.

There are (at least) two classes of intransitive verbs. There are unergatives, whose single argument is base-generated pre-verbally (in [Spec,*v*P], like (8). These verbs tend to assign an agent theta-role. Then there are unaccusatives, whose single argument is base-generated post-verbally (as sister to V), like (9). Just like verbal passives, the argument is taken to be a direct object, which receives a theme theta-role, and then (in some languages) raises to subject position (either for case or EPP reasons).

(8) The man jogged.

$[_{vP}$ $[_{NP}$ the man$]$ $[_{VP}$ jogged$]]$

$[_{IP}$ $[_{NP}$ the man$]_i$ $[_{vP}$ t_i $[_{VP}$ jogged$]]]$

(9) The ship sank.

$[_{VP}$ $[_{V'}$ sank $[_{NP}$ the ship$]]$

$[_{IP}\ [_{NP}\ \text{the ship}]_i\ [_{VP}\ [_{V'}\ \text{sank}\ t_i]$

Evidence for the unaccusativity analysis (object-to-subject movement) comes from several sources. As discussed in Section 1.2.1, resultative phrases (RPs) are restricted to modifying direct objects (Levin and Rappaport Hovav, 1995). So in (10), the RP can only be interpreted as meaning the plate became clean as a result of Mary's licking it, not that Mary herself became clean.

(10) Resultative Phrases

a. Mary licked the plate$_1$ clean$_1$.

b. *Mary$_1$ licked the plate clean$_1$.

As in the case above with the subject of a transitive verb, RPs are also not licit with subjects of unergatives (11).

(11) *Mary ran tired. (ungrammatical on RP reading)

RPs, however, *can* appear with unaccusatives (12). This is easily accommodated if the unaccusative argument is assumed to be a direct object that later moves.

(12) The river froze solid.

Further evidence for the movement analysis in unaccusatives comes from *Ne*-cliticization in Italian. *Ne* is a clitic that replaces partitive complements of nouns. It can only be extracted from direct object positions, and attaches to INFL (Belletti and Rizzi, 1981). Thus, *ne* can associate with a direct object in a transitive sentence (13). It is, however, ungrammatical with an unergative subject (14). Unlike the case of the unergative sentence, *ne* is fine with an unaccusative (15). Once again, this fact is naturally captured if it is assumed that the unaccusative argument is base-generated in object position where it licenses the partitive clitic.

(13)	Giovanni	ne	ha	insultati	due
	John	of-them	has	insulted	two
	'John has insulted two of them.'				

(14)	*Ne	telefonano	molti
	of-them	telephone	many
	'Many of them telephoned.'		

(15)	Ne	arrivano	molti
	of-them	arrive	many
	'Many of them arrived.'		

In many respects, then, unaccusatives and verbal passives are quite similar syntactically. Both involve direct objects that *A*-move to subject position. What then do the grammatical theories from Section 1.2.6 have to say about unaccusative acquisition? ACDH and EARH predicts that unaccusatives should be ungrammatical for premature children, on the reasonable assumption that unaccusatives do not project an external argument. CAH also predicts ungrammaticality for unaccusatives, since their theme theta-role in subject position is non-canonical (relative to unergatives). UFH, however, predicts no problem for unaccusatives, since their adult derivation does not involve movement from an already moved phrase, which would cause problems since that sort of movement requires an exception to Freezing. Finally, UPR expects problems for unaccusatives. Since they involve movement from the complement of the *v*P, unaccusative *v* must be defective (weak) in adult grammar. This *v*, however, will be phasal for premature children on UPR, and thus, movement to subject position will be blocked by PIC. What then is known about young children's comprehension of unaccusatives?

It happens that the empirical data are actually rather mixed. Before turning to these conflicting data, however, it is worth considering what sorts of errors one might expect to find if indeed unaccusatives are ungrammatical in early child grammar. Guasti (2002) provides three hypotheses about such a possibility. First, young (premature) children might simply avoid all unaccusative verbs. As a quick search of any child speech database will make clear, though, children do not avoid unaccusatives. There are hundreds of examples of unaccusative verbs in child-produced speech, from the earliest ages. Second, it might be that children will fail to raise the argument from object to subject position. That is, based on Baker's (1988) Uniformity of Theta Assignment Hypothesis (UTAH; which almost certainly has to be part of UG, and very likely required for verb learning in acquisition), children know that a theme must be generated in object position, but perhaps due to some grammatical restriction (e.g. CAH, UPR, etc), children fail to move the argument. This would of course violate other grammatical operations (e.g. EPP, case assignment), but might be the best children can do. While Déprez and Pierce (1993) do report finding some cases of post-verbal arguments with unaccusatives in early child productions (but never post-verbal arguments with unergatives), Stromswold (1996) reports such cases to be the exception rather than the rule. Most child productions with unaccusatives have pre-verbal subjects. Finally, children might simply treat unaccusatives as unergatives, since the latter are not predicted to be problematic on any of the grammatical acquisition theories (with the possible exception of ACDH given VPISH). That is, unergatives might serve as a s-homophone for unaccusatives (Babyonyshev, Ganger, Pesetsky, and Wexler, 2001). If this were the case, there are important consequences for children's unaccusatives. It begs the question of what theta-role children would associate with unaccusatives. If a theme, then would this constitute a violation of UTAH? If children's unaccusatives are unergatives, does this place an animacy restriction on the arguments that may appear with such verbs? These questions have yet to be addressed in the literature. What unaccusative data do exist, however, are reviewed next.[9]

Several studies report finding evidence for the early acquisition of unaccusatives, which would be in accord with UFH, but not UPR. Guasti (2002), examines a young (two-year-old) Italian child's choice of auxiliaries with unaccusatives. In Italian (and other Romance languages), unaccusatives select *be* and unergatives select *have*. If

---

[9] An exhaustive consideration of all acquisition studies concerning children's early knowledge of unaccusatives is simply beyond the scope of this dissertation. The most influential papers are reviewed here, but some important work is omitted.

children analyze unaccusatives as unergatives, one might expect all child productions with unaccusatives and an auxiliary to be of the (unergative) *have* type. Of the 22 sentences with an auxiliary and unaccusative verb that were found, however, only three included the wrong auxiliary, the rest correctly had *be*. Guasti takes this as evidence of early unaccusative acquisition. Ultimately, however, such data are not that persuasive. It could easily be the case that even in light of a grammatical deficit, children are nonetheless able to work out from the adult child-directed speech which auxiliary goes with which verb. Such a mapping is not dependent on having a deeper syntactic understanding of why unaccusatives select *be* and why unergatives select *be*.

A more compelling case for early unaccusative acquisition comes from children's use of reflexive clitics in Romance languages. Evidence exists that Romance reflexive object clitics (16), but not non-reflexive object clitics (17), involve unaccusative syntax (e.g. Marantz, 1984).[10] On such an account the reflexive clitic, unlike the non-reflexive clitic, is not a direct object. Rather, the surface subject in the reflexive sentences is actually an underlying object.

(16) Je me suis lavé *t*.

I me am washed

'I washed myself.'

(17) Je l' ai lavé.

I it have washed

'I washed it.'

[10] Examples taken from Snyder and Hyams (2005).

The reflexive clitic in (16) appears with the auxiliary *be*, while the non-reflexive clitic in (17) appears with the auxiliary *have*. The reflexive clitic appears with the same auxiliary (i.e. *be*) used with (most) unaccusative verbs (18).

(18) Je suis parti *t*.
I am left
'I left.'

Snyder, Hyams, and Crisma (1995) examined productions from a few French and Italian children (four total).[11] Of the 221 utterances they found with reflexive or non-reflexive clitics, only three auxiliary selection errors were noted (which all involved using *be* with a non-reflexive). The authors take appropriate auxiliary selection with reflexive and non-reflexive clitics as evidence that young children comprehend unaccusatives. These data are stronger than those presented by Guasti because here auxiliary selection is not simply a reflex of the main verb used, but of the type of clitic that appears with the main verb. If reflexive clitics do indeed involve unaccusative syntax, their early acquisition is problematic for UPR, but in line with predictions from UFH (but see Reinhart and Siloni, 2004 for arguments against an unaccusative analysis of reflexive clitics).

Evidence for delayed unaccusative acquisition comes from several studies. The first to be considered deal with case marking omissions in Japanese and Korean. Japanese overtly case marks subjects (*-ga*) and objects (*-o*). The accusative marker *–o* is usually omitted in adult (and thus child) language, presumably for phonological spell-out

[11] The data that follow are basically reconfirmed in Hyams and Snyder (2005).

reasons. In child produced speech, the NOM(inative) case marker almost always appears on subjects for unergatives and transitives. With unaccusatives, however, it is often found to be "dropped" or missing. Miyamoto, Wexler, Aikawa, and Miyagawa (1999) report that Japanese children omit 63% NOM on subject with unaccusatives, while only omitting 23% NOM with unergatives and transitives. Ito and Wexler (2002), in an even more careful test (which ruled out other reasons for case dropping), found omission of NOM on the subject of unaccusative to be 40%, and for unergatives and transitives, omission to be only 8%. Similar findings obtain for Korean. Lee and Wexler (2001) find that for Korean children, omission of NOM on subjects of unaccusatives is 65%, while for unergatives and transitives it is only 28.3% (Korean has the same properties as Japanese in this regard, i.e. empty phonetic spell-out of object case markings and SOV word order).

These results have been interpreted as demonstrating that young Japanese and Korean children fail to raise the object to subject position, whereby it receives object case, which can be dropped in accordance with adult speech.[12] While unaccusative arguments in early child English, as noted earlier, do not tend to appear in object position, such failures to raise might be more amenable in SOV languages since no word order violation is derived. That is, in an SOV language, the unraised structure is an s-homophone for the adult structure, whereas in an SVO language, the unraised structure is not an s-homophone for the adult structure, since only in the latter case is a word order violation noted.

[12] By what mechanism the object receives accusative case is left open. Presumably the <u>un</u>accusative verb is unable to assign (accusative) case. Perhaps accusative case is assigned as a default.

Recent evidence for unaccusative delay in English comes from work by Hirsch and Hartman (2006b). As part of a larger study into children's comprehension of experiencer passives, they tested comprehension of agentive (19) and nonagentive experiencer actives (20). Children correctly (100% accuracy) comprehend active sentences with object experiencer verbs (e.g. *scare*) when the subject is agentive. For example, in a story where a witch purposely jumps out from behind a tree and says "Boo!" to a young girl (Mary), children accept (19) as being true (Figure 4.1.4).

(19) The witch scared Mary.

(20) The spider scared Mary.

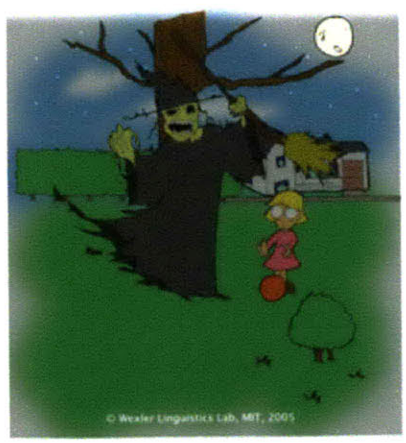


Figure 4.1.4: Picture used to depict agentive *scare* in Hirsch and Hartman (2006b).

The same verb *scare* used with an inanimate (nonagentive; dead spider) subject, however, causes many of children difficulties (only 59% accuracy).[13] Thus, in a story

[13] It is very interesting to note that the same children have no difficultly comprehending the passive of *scare*, in both its agentive and nonagentive uses. In fact, the passives with object-experiencer verbs are

where Mary walks into a room and notices a dead spider on a rock, which results in her screaming since she hates spiders, children incorrectly judge (20) to be false (Figure 4.1.5).

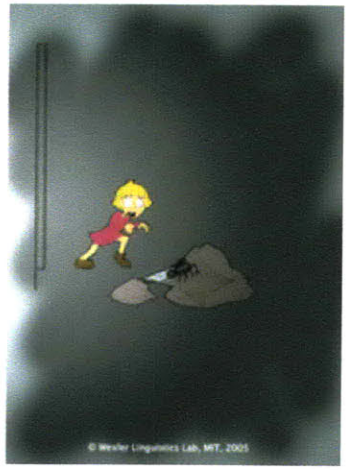


Figure 4.1.5: Picture used to depict nonagentive *scare* in Hirsch and Hartman (2006b).

Children's responses clearly demonstrate that they are unable to access the nonagentive interpretation even though the story makes clear that is what is called for, but instead children access the agentive interpretation only (Table 4.1.1.).

"Because the spider's not doing anything."	Camille, 3.14 years
"Cause he's [pointing to spider] not doing anything."	Marie, 3.91 years
"The spider's just sitting there."	Sebastian, 4.14 years
"Because the spider's not scaring him. He's just sitting there."	Alexis, 5.01 years
"Because the spider's not doing nothing."	Jailine, 5.85 years

actually comprehended earlier than passives with paradigmatic actional verbs such as *push* and *hit* (Hirsch and Hartman, 2006b). This is understandable given the adjectival strategy, since *scare* makes a better adjectival passive than either *push* or *hit*.

Table 4.1.1: Example responses to the question: "Why was the puppet wrong?" to nonagentive *scare* sentences in Hirsch and Hartman (2006b)

Belletti and Rizzi (1988) provide evidence from Italian that the underlying representation for nonagentive (but not agentive) *scare* (their *preoccupare*-class) involves movement of the Theme from an object position to subject position. If nonagentive object experiencer verbs involve unaccusative syntax as Belletti and Rizzi argue, then these acquisition results are predicted by UPR (but not UFH).

Finally, evidence for early child difficulties with unaccusatives comes from a Russian acquisition study by Babyonyshev, Ganger, Pesetsky, and Wexler (2001). In Russian, a direct object within the scope of negation can be realized with accusative case (if the object is definite/specific) or with genitive case (if the object is indefinite/non-specific). This Genitive of Negation (GoN) cannot apply to subjects (including those of transitives or unergatives). That is, the ability to be marked with genitive case is a property only of VP-internal indefinite objects (Pesetsky, 1982). In Russian, arguments of passives and unaccusatives can be realized in post-verbal, VP-internal position. They, too, can be marked by GoN. A small class of verbs, including existential *be*, requires its argument to be marked by GoN regardless of definiteness.

Even though arguments of passives and unaccusatives are realized in VP-internal position in Russian, there is strong evidence that they move covertly to subject position.[14]

[14] See Babyonyshev et al (2001) for details. Basically, evidence comes from the following details concerning the licensing of negative phrases. First, clausal negation is licit with negative phrases in the same clause. Second, negation in a lower clause cannot license a negative phrase in a higher clause. Third, a raised negative phrase subject must raise to a clause with negation, not from a clause with negation. Fourth, and crucially, a raising verb embedding a clause with an unaccusative and a genitive negative phrase must have negation in the higher clause, not down in the lower clause. That is, the last point only makes sense given the first three if the unaccusative argument has raised to the higher clause. The fact that

The strength of the GoN construction, since it requires the unaccusative argument to be pronounced in its base generated position, is that it rules out an unergative analysis (again, since unergatives do not allow genitive case). Without a s-homophone available, and to the extent that grammatical acquisition theories treat overt and covert movement in the same manner, UPR predicts children should not be able to mark genitive case with unaccusatives.

In order to test for the acquisition of GoN, a multi-part experimental investigation was needed. Native Russian children from 3;0 to 6;6 years old were tested on various structures. First, it must be established if and which children have acquired GoN, independent of its use with unaccusatives. To do so, children's use of GoN with transitives was tested. The finding across all children is that they correctly use GoN about 75% of the time where they should, and only apply it about 4% of the time where they should not. The general finding from unaccusatives is that for both those verbs that require GoN for all arguments (e.g. *be*) and those that require it just for indefinite objects, arguments are much more rarely marked with genitive than in transitives. There is a strong age effect as well. Younger children (mean age 4;0) use GoN 30% of the time, while older children (5;4) use it 60% of the time. Looking at the data from individual children is particularly revealing. Some children do not use GoN correctly under any circumstances (even with transitives). Clearly these children's data therefore cannot speak to the issue of unaccusative acquisition using GoN. Some children use GoN correctly in all cases. These are the children who are adult-like in their performance and can be taken as those who have matured. Finally, the rest of the children use GoN

---

the object does not appear overtly in the higher clause suggests that this movement takes place covertly (e.g. at LF).

correctly with transitives, but not unaccusatives (even with *be* which always requires genitive case marking under negation). Crucially, there are no children who use GoN correctly for unaccusatives, but not for transitives.

All together then, the results from research into early child use and comprehension of unaccusatives is mixed. Studies examining early child use of reflexive clitics suggest unaccusatives are acquired early. Evidence from case marking omissions, active nonagentive object experiencer verbs, and GoN studies all point to either non-adult analyses or comprehension delays with unaccusatives. While the latter studies are supportive of UPR, the former are not (and argue for something more akin to UFH). Ultimately, there is a need for (a) further unaccusative acquisition studies, especially cross-linguistic investigations given generally poor unaccusativity diagnostics in English, and (b) within-subject tests of passives/raising and unaccusatives. The latter would help conclusively demonstrate whether or not the grammatical issue that causes passives and StS raising sentences to be delayed is manifested in unaccusatives. We now turn away from unaccusatives and examine how UPR, which predicts StS raising to be delayed, can account for the early acquisition of control sentences if control involves raising.

Hornstein (1999) argues that Minimalist considerations lead us to eliminate the control module, and that obligatory control is best understood as a case of syntactic raising. There is controversy in the syntactic literature on this point: Culicover and Jackendoff (2001) as well as Landau (2003) argue against Hornstein's proposal, and Boeckx and Hornstein (2003) reply. A complete review of this debate is beyond what can reasonably be considered in this space, but an attempt will be made to see how Hornstein's proposal meshes with the developmental results obtained in Studies 1-7 and what is known about the acquisition of control (for review, see Section 3.3.1). What does control as raising predict about development and how do these predictions fit the facts?

Initially, one would think that if control is raising, then the prediction is that control and raising should develop at about the same time and involve similar developmental trajectories. Since this is greatly at odds with the acquisition facts (see Section 3.3.1 for a review of evidence that [obligatory] control is very early in language development), one might be inclined to conclude that the developmental data do not support the control as raising hypothesis.

But let us do a bit more justice to this hypothesis. Let us look at the proposal in detail to see if we can be more explicit about the relation between the proposal and the developmental principles (like UPR) that are known. Hornstein's (1999) analysis is illustrated in (21) (his (19)):

(21) John hopes to leave

[IP John [VP John [hopes [IP John to [VP John leave]]]]]

Hornstein's explanation of the derivation in (21) is as follows. *John* merges with *leave*. *John* then raises to the embedded [Spec, IP]. *John* raises again to [Spec, VP] of *hope*. By principles that Hornstein adumbrates, the chain that *John* heads has "two theta-roles, the leaver role and the hoper role." *John* then raises to [Spec, IP] of the matrix clause.

Clearly there are *A*-chains here, but ACDH is not under discussion, UPR is. Is UPR violated by this analysis? It is hard to tell because Hornstein does not discuss an analysis incorporating phasal considerations, that is, any kind of strong cyclicity. There is no *v*P, only VP. But let us see what seems reasonable if we attempted to understand the derivation in phasal terms. Suppose there is a phasal head *v* between the embedded IP and the lower VP. The first raising, from [Spec, VP] to [Spec, IP] would be the movement of a phrase in the complement of *v* (after all, it is in the VP that *v* selects) to the next higher phase. If *John* is actually merged into [Spec, *v*P], then raising to [Spec, IP] is licit given PIC because *John* comes from the edge of the next lower phase. So far no phasal violations are incurred on the adult analysis, similarly for the child analysis, since no defective phases are involved.

Now *John* raises to [Spec, VP] of *hope*. Let us assume again that *hope* is introduced in a VP that is a complement of a *v*. This *v* is phasal, since *hope* assigns an external argument. Since the raising of *John* allows it to check the external theta-role feature of *hope*, let us assume that the raising goes to [Spec, *v*P].[15] Then at phase CP, T can attract *John*, in the edge of the lower phase, and it can raise for the last time. Under this analysis it looks as if no defective phase is needed in the derivation. The strict cycle can be followed, with all material except edges shipped off to interpretation and not available at the next stage, except the edge of the lower phase. Thus for the child, UPR is not violated. The child will be able to compute the (raising) derivation for control.

If there is any reason that the first raising of *John* must go to [Spec, VP] of *hope* rather then to [Spec, *v*P] then we are in a different situation. Since *John* is then in the complement of the highest *v*, T cannot attract it, and the highest *v* will have to be non-phasal. This seems to be against the spirit of phase theory, since *hope* assigns an external

---

[15] Chomsky's system does not allow movement to the first [Spec, *v*P], but presumably this could be allowed in another system. Details would have to be worked out.

argument feature. At any rate, if this were the case in the derivation, then there would be a UPR situation for the child; the child would take the highest *v* as phasal and the derivation would crash, predicting difficulties for control structures like (21).

But it seems most reasonable, until further considerations come in, to take the former analysis, with movement to [Spec, *v*P]. If this is indeed correct, then UPR predicts no trouble for control, even with a raising analysis of control. In such a case, the development of control cannot distinguish the control module versus raising analysis, at least not in these terms. Raising, presumably, would still look the way it traditionally does, a defective *v* will be needed, and raising will be predicted to be late.

It is quite interesting, then, that UPR, a theory of acquisition that relates to phases rather than to chains, seems to predict no problems for children on control, even with a raising analysis. The chains are not what matter; only the phases and their conditions (PIC) matter. This should be a familiar lesson: labels ('raising') are not as important as analyses. On deeper inspection, even with a raising analysis, control might not demand extraordinary conditions on movement; raising does, passives do, and unaccusatives do.

We make these observations quite tentatively; deeper syntactic analysis might contradict us, but it will take that. For the moment, we predict good control and poor raising until UPR matures, whether control is a control module or a raising analysis of the sort that Hornstein provides. These predictions agree with the developmental data on control which, as outlined in Section 3.3.1, is acquired early.[16]

[16] Boeckx and Hornstein (2003) argue that the late development of subject control with verbs like *promise* (when taking an 'object'; e.g. *A promised B to do Z*) supports the movement analysis of control because object control is expected given the Minimal Link Condition and the movement analysis. Thus subject control is "marked" and late development is expected. This analysis, however, does not account for the extremely late development of subject-to-subject raising. Why should it take such a huge amount of time for subject-to-subject raising to develop given that it is possible and is in the adult input? Further, as Boeckx and Hornstein themselves note based on informal observations, *promise* with an object and non-

4.1.3 *UPR and Linguistic Theory*

The acquisition timeline of various grammatical structures (e.g. verbal and adjectival passives, unraised and raised sentences with StS raising verbs, with and without an experiencer-phrase, and unaccusatives) have now been considered, most, if not all of which comport with the predictions of UPR. It would therefore appear that UPR is a strong candidate for explaining how one aspect of adult grammar comes about in development. It is the case, however, that UPR fundamentally relies on a fair number of rather specific assumptions about the nature of human grammar, most of which are reiterated below.

Chomsky (2001) argues that syntactic derivations proceed incrementally, one phase at a time. Each phase is made up from a unique lexical array ("numeration"), and consists of the head of the phase (where only *v* and C are taken to define phases), the complement (sister) of the phase, and the edge (specifier) of the phase. Once a phase is built, most of its content (the phase complement) is transferred to the interfaces and is no longer relevant (accessible) for subsequent higher syntactic derivation. Only the phase edge and head remain accessible to the next higher phase (this is Chomsky's Phase Impenetrability Condition). Only by moving to the phase edge can material in the phase

---

finite complement is ungrammatical for many English-speakers. If this is true, which we think it is based on some of our own recent pilot studies, then poor comprehension by many of the children cannot be taken as evidence for late acquisition, but merely as a reflection that such structures are ungrammatical for many speakers. Also, it bears noting that the relevant structure (*A promised B to do Z*) never appears in the child-directed speech for any (of 1051) English-speaking children on the CHILDES corpus. That is, children are never presented evidence in the form of adult speech that *promise* is a possible control verb when an object is also present. Wexler (2004) suggests that Larson's (1991) account, which involves an A-chain (and presumably defective *v*) in the analysis of *promise* subject control, together with ACDH (or UPR) predicts the late development of such structures. A reasonable research strategy would be to see if the correct analysis of *promise* involves a non-phasal *v* so that UPR would predict that the computation would not converge for the premature child.

complement become available for future movement higher. Together, these principles give rise to a form of successive cyclic movement, whereby all long-distance movement must occur through successive phase edges. In order to account for movement that appears to cross a phase boundary (e.g. verbal passives, raised sentences, and unaccusatives), Chomsky introduces the distinction of weak and strong phases. PIC only applies to strong phases, where Agreement and movement from a higher phase into the complement of a lower phase is only possible when the lower phase is weak (i.e. defective).

In particular, UPR fundamentally rests upon the validity in adult grammar of strong cylicity, phases (both weak/defective and strong/non-defective), PIC, and lastly, that the Smuggling analysis of raising is not correct. If any of these grammatical assumptions were shown not to hold, UPR would lose its explanatory power. In this section, we briefly consider the validity of the syntactic framework required by UPR.

A first question is: are phases psychologically real? A related question is: what are the phases, and what evidence do we have for their existence; in particular, does a strong/weak dichotomy exist? Chomsky (2001) argues for the existence of phases on both conceptual and empirical grounds. Conceptually, phases help significantly reduce total required computational resources. Chomsky also argues that they are defined substantively on an argument from completeness of semantic properties. This last point, however, is somewhat dubious. It is unclear, for example, how the *v*P is semantically "complete" if later movement can strip it of the external argument. Legate (2003) offers several pieces of evidence against Chomsky's strong/weak phase partition. She shows that passive and unaccusative *v* patterns with transitive *v** in cases of reconstruction, Antecedent Contained Deletion, and parasitic gap construction. If her arguments are valid, then UPR is without a mechanism for distinguishing between which structures are acquired early and which are acquired late. Dikken (2006), however, argues that all of Legate's arguments are inconclusive on grounds that they are either based on invalid assumptions or are inconsistent with a broader array of data. Boeckx and Grohmann (2007) survey developments in linguistic theory relative to phases, and conclude that while there exist many challenges to Chomsky's model, the intuition is sound and useful, and no better approach is currently available.

Collins' (Smuggling) analysis of verbal passives is compatible with both UFH and UPR. That is, if verbal passives do involve Smuggling, UPR would still predict passives to be delayed (whether they contain a by-phrase or not), since his analysis requires movement from the complement of *v* (which UPR rules to be ungrammatical for young children). Of crucial relevance, however, is that if his Smuggling account of raising over an experiencer is correct, UPR predicts no delay for ROE sentences. Collins' analysis never involves movement from the complement of *v*. In fact, for Smuggling to work, it involves the embedded subject (as part of a VP) moving to [Spec,*v*P], thus able to move to matrix T since PIC is not violated.

There is absolutely no question, though, that Collins' analyses make many theoretical assumptions that receive limited support. An account is needed of Freezing Principle "exceptions". Also, for example, both the early movement of the embedded subject to [Spec,VP] (bringing the embedded subject past *seem*, though the two never enter an Agree relation) and the later Smuggling movement (bringing the VP containing the embedded subject, past the experiencer, to spec,*v*P) could use further independent motivation and evidence. Finally, sentences with raising and *wh*-movement appear to be a challenge. If the VP in the Smuggling step moves to [Spec,*v*P], how can *who* raise if it too must occupy that position at some point (22)?

(22) Who does John seems to Mary to love?

Ultimately, the validity of the Smuggling approach rests on the soundness of its assumptions and whether it can be extended to languages other than English. Whether Collins' analysis can be extended to cover the full range of crosslinguistic phenomena has yet to be made clear. It is, however, worth repeating, that *if* Smuggling, as exactly envisioned by Collins, is the correct syntactic derivation for ROE, *then* UPR does not predict difficulties for premature children, since no movements take place out of the complement of any *v*, which is decidedly different than on Kim's (2005) account, and on which UFH would not predict difficulties for premature children.

A final issue briefly worth mentioning is that the relevance of the RNE sentences in distinguishing between UFH and UPR would be completely undermined if the tested RNE sentences involve a covert experiencer. If they do, then UFH could rule them out on the assumption that they involve the Smuggling operation that UFH assumes for ROE sentences. It is unclear if there are any syntactic tests that can diagnose a possible covert experiencer. If English RNE sentences with *seem* and *appear* contain a covert experiencer, the data obtained from StS raising in Studies 4 and 5 would be equally compatible with UFH and UPR, and one would have to turn either to other structures (e.g. unaccusatives) or languages in which ROE is ungrammatical (and in which RNE sentences can therefore perhaps more safely be assumed not to involve covert experiencers) in order to decide between them.

4.1.4 *Conclusion*

Production data from Study 7 demonstrate that while StS raising is not uncommon in child-directed speech, children very rarely use such structures in their own productions. This finding argues for grammatical acquisition theories positing StS raising to be delayed in child grammar, and suggests that the compensatory heuristics discussed in Chapters 2 and 3 are likely products of children seeking real-time interpretations for otherwise ungrammatical strings.

Consideration of how and when unaccusatives are acquired failed to offer definitive support for either UPR or UFH. Some studies suggest unaccusatives are early in acquisition, while others point to delays. Future unaccusative studies are needed to draw any strong conclusions about how this structure can inform grammatical acquisition research. Control as Raising was also considered in brief, and at least at first blush, appears to be compatible with grammatical accounts that rule out StS raising.

Finally, issues in theoretical (adult) syntax were considered. A great deal of syntactic machinery, much of it on somewhat debatable empirical footing, is required for the predictions of UPR to go through to account for verbal passive and StS raising delay. Of course, the same is true of UFH. The adequacy of UPR as an account of these acquisition phenomena is ultimately dependent on a specific syntactic framework being validated.

In the next section, some final issues as relate to StS raising in biolinguistics are considered.

4.2 Summary and Future Directions

The work presented in this dissertation offers the first comprehensive investigation in the literature into children's acquisition of subject-to-subject raising. This research was undertaken both to offer a look at a previously ignored area of acquisition, one that is important in (adult) syntactic theory, and as a means of further studying the validity of several grammatical accounts of verbal passive acquisition. The seven studies included here strongly argue that StS raising is late in language development, and the grammatical reason for this delay is related to that governing verbal passives. One can now ask where this work should lead. Three broad areas for continued research are outlined: issues specific to the acquisition of StS raising, general issues in language acquisition research, and implications for biolinguistics.

The StS raising studies undertaken here sacrificed breadth for depth of inquiry. Ultimately, the studies here are limited to a very small set of StS raising verbs (*seem* and *appear*), mostly involving only comprehension studies, and apply only to English-speaking children. Hopefully future studies will expand on all three aspects (e.g. for recent results on the acquisition of StS raising in Dutch, see Koring and Wexler, 2009). First, more StS raising verbs must be considered, in particular those that do not allow an experiencer-phrase (see Section 4.1.3). Second, comprehension should be investigated for

StS raising adjectives (e.g. *likely*, *certain*) to see if predictions apply outside of the verbal domain. Third, other raising structures should be tested (e.g. modal auxiliaries). Fourth, crosslinguistic studies are needed to test for the (predicted) universality of StS raising delay. Fifth, more work in the area of production studies ought to be undertaken to add to the comprehension results.

The validity of particular grammatical acquisition theories is dependent on their ability to make correct predictions about *all* the data they cover. To the extent that the various accounts in Section 1.2.6 all overlapped in their correct predictions for verbal passives acquisition, it was readily apparent that one had to seek comprehension data for all structures on which those accounts would make different predictions. StS raising is one such structure, but there are many more. A more thorough review of the literature (plus future experimental acquisition studies) will inevitably lead to a clearer picture of just what aspects of grammar are difficult for children, and how best to cognitively model any such findings. For example, while UPR, of the acquisition theories considered here, is most compatible with the StS raising data from Studies 1-7, it is not yet clear if UPR is tenable given all the evidence from unaccusative acquisition (Section 4.1.2). That is, if UPR fails to predict early unaccusative acquisition (or at least earlier than verbal passive and StS raising), then it would fare about as well as UFH, which only had trouble predicting delays for RNE sentences (Studies 4-6). To better help tease apart UPR and UFH, and perhaps even build a new grammatical acquisition theory, the acquisition of other structures must be considered. Three that hold some interesting promise, as they involve long-distance movement and have been argued to involve Smuggling, include:

tough-movement (Hicks, 2009), locative inversion (Rizzi and Shlonsky, 2006), and clefts and inverse copulas (Hirsch and Wexler, 2007b).

Above and beyond the interesting questions concerning when StS raising emerges in child development, or how best to grammatically capture any noted delays, there exist many other significant and important questions about how this work fits into the larger picture of biolinguistics. One compelling area to which it contributes is the understanding of what role biological maturation plays in language acquisition. While extensive evidence was marshaled in Section 1.2.5 in support of verbal passives being subject to maturation, through developmental association with verbal passives (Studies 3 and 6), StS raising also seems a likely candidate for maturation. More studies, however, are needed. The type of behavioral genetics study that Ganger, Dunn, and Gordon (2004) conducted for passives could be extended to StS raising, as could the environmental variables study by Hirsch, Modyanova, and Wexler (2006).

More ambitious studies should also be considered, linking grammatical deficits and their biological causes. By examining correlations between brain development and linguistic development, neural areas underlying language representation and processing might be discovered. With putative neural language areas in mind, neuroimaging studies could be conducted in adults to further investigate what roles those areas might play in language function. Likewise, to the extent that grammatical theories predict that certain grammatical operations and structures share similarities, neuroimaging studies ought to find common neural networks of activation in such cases.

With a more accurate understanding of how language typically develops comes the possibility for research either into how development goes awry, or how acquired language functions degenerate (either due to acute brain damage or disease). Having gained a better understanding of the time course and details of StS raising acquisition in typically developing children, we can now turn our attention to how this grammatical structure is comprehended in other (atypical) populations. Some interesting areas (already known for language related deficits) for future study include Specific Language Impairment (SLI), Williams syndrome, Down syndrome, autism, Broca's aphasia, and Parkinson's disease. To the extent that knowledge of StS raising is impaired in any of these populations, comparing brain areas that are affected in each case might lead to a better understanding of how StS raising is instantiated and processed neurally. Ultimately, having a fuller understanding of language impairments in the above cases might eventually lead to positive clinical applications. Studies into how StS raising is learned in second language acquisition would also be of interest.

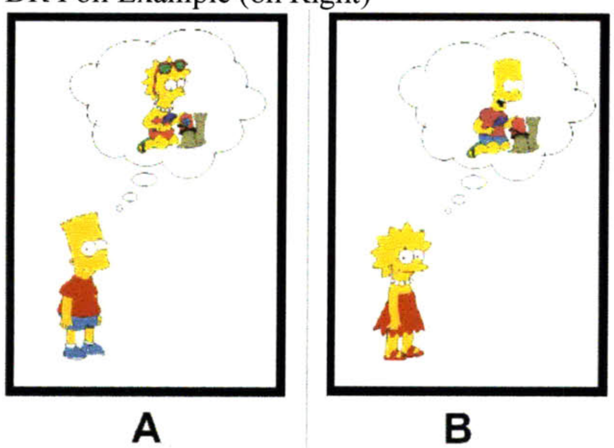
Given ever-continued improvements in linguistic theory, novel experimental approaches and greater breadth of coverage in language acquisition studies, increasing access to patient populations, and advances in the fields of neuroscience and genetics, this is an extremely exciting time to be involved in research concerning biolinguistics. Or, at least, so it would *seem*.

Appendices

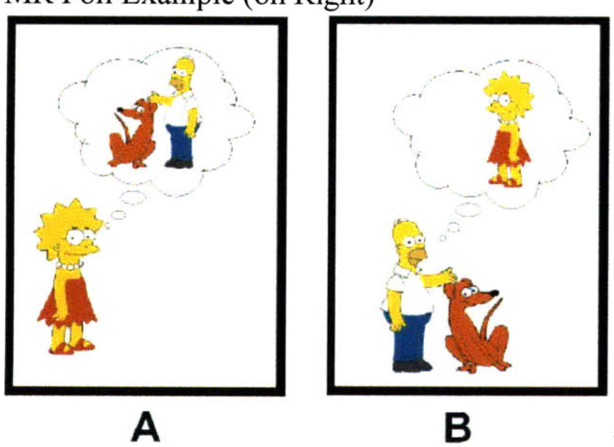
A1: Study 1

Example Pictures:

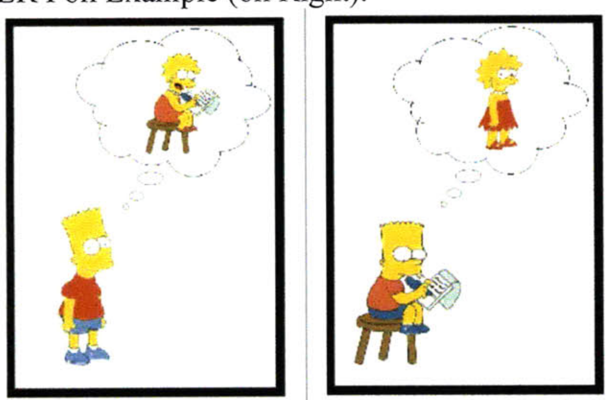
DR Foil Example (on Right)



MR Foil Example (on Right)

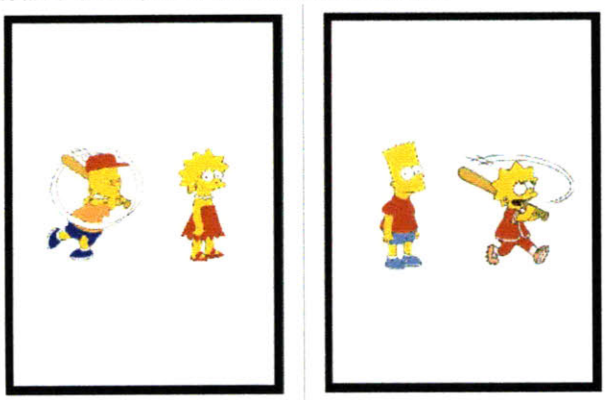


ER Foil Example (on Right):



A B

Active Transitive Condition Picture:



A B

Test Items (as Linger file).

Guide to test items (as Linger file; for details see http://tedlab.mit.edu/~dr/Linger/)
Line 1: # <study name> <item number> <condition>
Line 2: <image file>
Line 3: <test sentence>
Line 4: <side of screen on which correct choice appears>

Test items:

Study-1 1 R-DR
Images/1.jpg
Lisa seems to Homer to be rolling a ball.
R

Study-1 2 R-DR
Images/2.jpg
Homer seems to Marge to be eating a meal.

L

Study-1 3 R-DR
Images/3.jpg
Bart seems to Marge to be waving a flag.
R

Study-1 4 R-DR
Images/4.jpg
Lisa seems to Bart to be building a sandcastle.
L

Study-1 5 R-DR
Images/5.jpg
Bart seems to Lisa to be kicking a ball.
L

Study-1 6 R-DR
Images/6.jpg
Marge seems to Homer to be licking a lamppost.
R

Study-1 7 U-DR
Images/7.jpg
It seems to Homer that Lisa is bowling a ball.
R

Study-1 8 U-DR
Images/8.jpg
It seems to Marge that Homer is eating a meal.
L

Study-1 9 U-DR
Images/9.jpg
It seems to Marge that Bart is waving a flag.
R

Study-1 10 U-DR
Images/10.jpg
It seems to Bart that Lisa is building a sandcastle.
L

Study-1 11 U-DR
Images/11.jpg
It seems to Bart that Lisa is kicking a ball.
R

Study-1 12 U-DR
Images/12.jpg
It seems to Marge that Homer is licking a lamppost.
L

Study-1 13 R-ER
Images/13.jpg
Marge seems to Homer to be reading a book.
L

Study-1 14 R-ER
Images/14.jpg
Homer seems to Marge to be driving a car.
R

Study-1 15 R-ER
Images/15.jpg
Lisa seems to Homer to be eating ice cream.
L

Study-1 16 R-ER
Images/16.jpg
Lisa seems to Bart to be playing a musical instrument.
R

Study-1 17 R-ER
Images/17.jpg
Bart seems to Lisa to be writing a letter.
R

Study-1 18 R-ER
Images/18.jpg
Bart seems to Marge to be digging a hole.
L

Study-1 19 U-ER
Images/19.jpg
It seems to Marge that Homer is reading a book.
R

Study-1 20 U-ER
Images/20.jpg
It seems to Marge that Homer is driving a car.
R

Study-1 21 U-ER
Images/21.jpg
It seems to Lisa that Homer is eating ice cream.
L

Study-1 22 U-ER
Images/22.jpg
It seems to Bart that Lisa is playing a musical instrument.
R

Study-1 23 U-ER
Images/23.jpg
It seems to Bart that Lisa is writing a letter.
L

Study-1 24 U-ER
Images/24.jpg
It seems to Marge that Bart is digging a hole.
L

Study-1 25 R-MR
Images/25.jpg
Lisa seems to Homer to be petting a dog.
R

Study-1 26 R-MR
Images/26.jpg
Bart seems to Lisa to be lifting a rock.
L

Study-1 27 R-MR
Images/27.jpg
Bart seems to Marge to be helping a cat.
R

Study-1 28 R-MR
Images/28.jpg
Lisa seems to Bart to be washing the floor.
L

Study-1 29 R-MR
Images/29.jpg
Homer seems to Marge to be pushing a cart.
L

Study-1 30 R-MR

Images/30.jpg
Marge seems to Homer to be splashing a lobster.
R

Study-1 31 U-MR
Images/31.jpg
It seems to Lisa that Homer is petting a dog.
L

Study-1 32 U-MR
Images/32.jpg
It seems to Bart that Lisa is lifting a rock.
R

Study-1 33 U-MR
Images/33.jpg
It seems to Marge that Bart is helping a cat.
L

Study-1 34 U-MR
Images/34.jpg
It seems to Bart that Lisa is washing the floor.
R

Study-1 35 U-MR
Images/35.jpg
It seems to Marge that Homer is pushing a cart.
L

Study-1 36 U-MR
Images/36.jpg
It seems to Marge that Homer is splashing a lobster.
R

Study-1 37 Tran
Images/37.jpg
Lisa is swinging a bat.
R

Study-1 38 Tran
Images/38.jpg
Bart is punching a bag.
L

Study-1 39 Tran
Images/39.jpg

Lisa is eating a sandwich.
L

Study-1 40 Tran
Images/40.jpg
Bart is throwing water.
R

Study-1 41 Tran
Images/41.jpg
Homer is playing with blocks.
R

Study-1 42 Tran
Images/42.jpg
Marge is walking a dog.
L

Study-1 43 Tran
Images/43.jpg
Lisa is swinging a bat.
L

Study-1 44 Tran
Images/44.jpg
Bart is punching a bag.
R

Study-1 45 Tran
Images/45.jpg
Homer is eating a sandwich.
R

Study-1 46 Tran
Images/46.jpg
Bart is throwing water.
R

Study-1 47 Tran
Images/47.jpg
Marge is playing with blocks.
L

Study-1 48 Tran
Images/48.jpg
Homer is walking a dog.

L

Study-1 49 Think
Images/49.jpg
Homer thinks Lisa is bowling a ball.
R

Study-1 50 Think
Images/50.jpg
Marge thinks Homer is eating a meal.
L

Study-1 51 Think
Images/51.jpg
Marge thinks Bart is waving a flag.
R

Study-1 52 Think
Images/52.jpg
Bart thinks Lisa is building a sandcastle.
R

Study-1 53 Think
Images/53.jpg
Bart thinks Lisa is kicking a ball.
L

Study-1 54 Think
Images/54.jpg
Marge thinks Homer is licking a lamppost.
R

Study-1 55 Think
Images/55.jpg
Marge thinks Homer is reading a book.
R

Study-1 56 Think
Images/56.jpg
Marge thinks Homer is driving a car.
R

Study-1 57 Think
Images/57.jpg
Lisa thinks Homer is eating ice cream.
L

Study-1 58 Think
Images/58.jpg
Bart thinks Lisa is playing a musical instrument.
L

Study-1 59 Think
Images/59.jpg
Bart thinks Lisa is writing a letter.
R

Study-1 60 Think
Images/60.jpg
Marge thinks Bart is digging a hole.
R

Study-1 61 Think
Images/61.jpg
Lisa thinks Homer is petting a dog.
L

Study-1 62 Think
Images/62.jpg
Bart thinks Lisa is lifting a rock.
R

Study-1 63 Think
Images/63.jpg
Marge thinks Bart is helping a cat.
L

Study-1 64 Think
Images/64.jpg
Bart thinks Lisa is washing the floor.
L

Study-1 65 Think
Images/65.jpg
Marge thinks Homer is pushing a cart.
R

Study-1 66 Think
Images/66.jpg
Marge thinks Homer is splashing a lobster.
L

A2: Study 2

Note: Same images used as in Study 1 (A1).

Guide to test items (as Linger file):
Line 1: # <study name> <item number> <condition>
Line 2: <image file>
Line 3: <test sentence>
Line 4: <side of screen on which correct choice appears>

Test items:

Study-2 1 R-DR
Images/1.jpg
To Homer, Lisa seems to be rolling a ball.
R

Study-2 2 R-DR
Images/2.jpg
To Marge, Homer seems to be eating a meal.
L

Study-2 3 R-DR
Images/3.jpg
To Marge, Bart seems to be waving a flag.
R

Study-2 4 R-DR
Images/4.jpg
To Bart, Lisa seems to be building a sandcastle.
L

Study-2 5 R-DR
Images/5.jpg
To Lisa, Bart seems to be kicking a ball.
L

Study-2 6 R-DR
Images/6.jpg
To Homer, Marge seems to be licking a lamppost.
R

Study-2 7 U-DR
Images/7.jpg
To Homer, it seems that Lisa is bowling a ball.
R

Study-2 8 U-DR
Images/8.jpg
To Marge, it seems that Homer is eating a meal.
L

Study-2 9 U-DR
Images/9.jpg
To Marge, it seems that Bart is waving a flag.
R

Study-2 10 U-DR
Images/10.jpg
To Bart, it seems that Lisa is building a sandcastle.
L

Study-2 11 U-DR
Images/11.jpg
To Bart, it seems that Lisa is kicking a ball.
R

Study-2 12 U-DR
Images/12.jpg
To Marge, it seems that Homer is licking a lamppost.
L

Study-2 13 R-ER
Images/13.jpg
To Homer, Marge seems to be reading a book.
L

Study-2 14 R-ER
Images/14.jpg
To Marge, Homer seems to be driving a car.
R

Study-2 15 R-ER
Images/15.jpg
To Homer, Lisa seems to be eating ice cream.
L

Study-2 16 R-ER
Images/16.jpg
To Bart, Lisa seems to be playing an instrument.
R

Study-2 17 R-ER
Images/17.jpg
To Lisa, Bart seems to be writing a letter.
R

Study-2 18 R-ER
Images/18.jpg
To Marge, Bart seems to be digging a hole.
L

Study-2 19 U-ER
Images/19.jpg
To Marge, it seems that Homer is reading a book.
R

Study-2 20 U-ER
Images/20.jpg
To Marge, it seems that Homer is driving a car.
R

Study-2 21 U-ER
Images/21.jpg
To Lisa, it seems that Homer is eating ice cream.
L

Study-2 22 U-ER
Images/22.jpg
To Bart, it seems that Lisa is playing an instrument.
R

Study-2 23 U-ER
Images/23.jpg
To Bart, it seems that Lisa is writing a letter.
L

Study-2 24 U-ER
Images/24.jpg
To Marge, it seems that Bart is digging a hole.
L

Study-2 25 R-MR
Images/25.jpg
To Homer, Lisa seems to be petting a dog.
R

Study-2 26 R-MR

Images/26.jpg
To Lisa, Bart seems to be lifting a rock.
L

Study-2 27 R-MR
Images/27.jpg
To Marge, Bart seems to be helping a cat.
R

Study-2 28 R-MR
Images/28.jpg
To Bart, Lisa seems to be washing the floor.
L

Study-2 29 R-MR
Images/35.jpg
To Marge, Homer seems to be pushing a cart.
L

Study-2 30 R-MR
Images/30.jpg
To Homer, Marge seems to be splashing a lobster.
R

Study-2 31 U-MR
Images/31.jpg
To Lisa, it seems that Homer is petting a dog.
L

Study-2 32 U-MR
Images/32.jpg
To Bart, it seems that Lisa is lifting a rock.
R

Study-2 33 U-MR
Images/33.jpg
To Marge, it seems that Bart is helping a cat.
L

Study-2 34 U-MR
Images/34.jpg
To Bart, it seems that Lisa is washing the floor.
R

Study-2 35 U-MR
Images/35.jpg

To Marge, it seems that Homer is pushing a cart.
L

Study-2 36 U-MR
Images/36.jpg
To Marge, it seems that Homer is splashing a lobster.
R

Study-2 37 Tran
Images/37.jpg
Lisa is swinging a bat.
R

Study-2 38 Tran
Images/38.jpg
Bart is punching a bag.
L

Study-2 39 Tran
Images/39.jpg
Lisa is eating a sandwich.
L

Study-2 40 Tran
Images/40.jpg
Bart is throwing water.
R

Study-2 41 Tran
Images/41.jpg
Homer is playing with blocks.
R

Study-2 42 Tran
Images/42.jpg
Marge is walking a dog.
L

Study-2 43 Tran
Images/43.jpg
Lisa is swinging a bat.
L

Study-2 44 Tran
Images/44.jpg
Bart is punching a bag.

R

Study-2 45 Tran
Images/45.jpg
Homer is eating a sandwich.
R

Study-2 46 Tran
Images/46.jpg
Bart is throwing water.
R

Study-2 47 Tran
Images/47.jpg
Marge is playing with blocks.
L

Study-2 48 Tran
Images/48.jpg
Homer is walking a dog.
L

Study-2 49 Think-DR
Images/49.jpg
Homer thinks Lisa is bowling a ball.
R

Study-2 50 Think-DR
Images/50.jpg
Marge thinks Homer is eating a meal.
L

Study-2 51 Think-DR
Images/51.jpg
Marge thinks Bart is waving a flag.
R

Study-2 52 Think-DR
Images/52.jpg
Bart thinks Lisa is building a sandcastle.
R

Study-2 53 Think-DR
Images/53.jpg
Bart thinks Lisa is kicking a ball.
L

Study-2 54 Think-DR
Images/54.jpg
Marge thinks Homer is licking a lamppost.
R

Study-2 55 Think-ER
Images/55.jpg
Marge thinks Homer is reading a book.
R

Study-2 56 Think-ER
Images/56.jpg
Marge thinks Homer is driving a car.
R

Study-2 57 Think-ER
Images/57.jpg
Lisa thinks Homer is eating ice cream.
L

Study-2 58 Think-ER
Images/58.jpg
Bart thinks Lisa is playing a instrument.
L

Study-2 59 Think-ER
Images/59.jpg
Bart thinks Lisa is writing a letter.
R

Study-2 60 Think-ER
Images/60.jpg
Marge thinks Bart is digging a hole.
R

Study-2 61 Think-MR
Images/61.jpg
Lisa thinks Homer is petting a dog.
L

Study-2 62 Think-MR
Images/62.jpg
Bart thinks Lisa is lifting a rock.
R

Study-2 63 Think-MR
Images/63.jpg
Marge thinks Bart is helping a cat.
L

Study-2 64 Think-MR
Images/64.jpg
Bart thinks Lisa is washing the floor.
L

Study-2 65 Think-MR
Images/65.jpg
Marge thinks Homer is pushing a cart.
R

Study-2 66 Think-MR
Images/66.jpg
Marge thinks Homer is splashing a lobster.
L

A3: Adult Judgments of *Expect*-Analysis Predictions

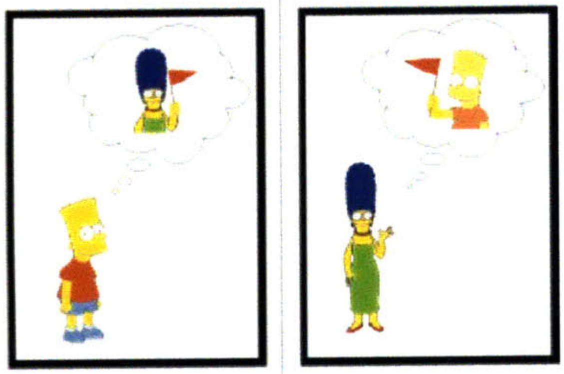
Note: Only native English-speaking adults (18 years old or older) were querried.

Instructions: "I could use your help with a VERY QUICK linguistics test for my dissertation research. All you have to do is open the attached PDF, and answer THREE questions. Note, there are NO CORRECT answers. I am simply interested in which picture you think BETTER / BEST matches the test sentence in each case (you must choose one picture for each sentence). To assist me, just REPLY to this e-mail, with your responses (A or B) filled out below."

File:

Question 1:

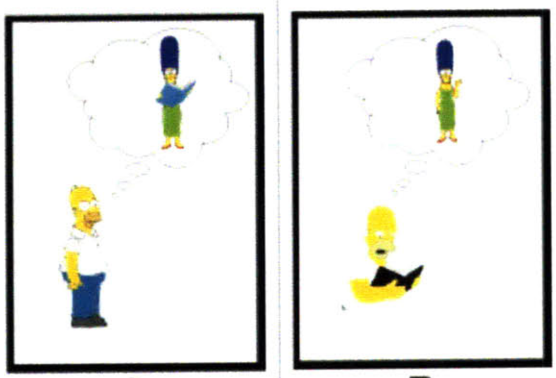
Which picture (A or B) *best* shows: **Bart expects to be waiving a flag.**



A B

Question 2:

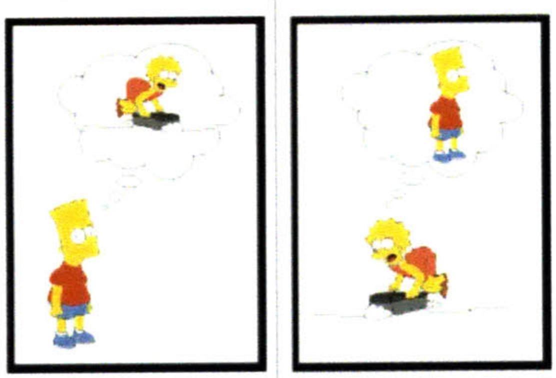
Which picture (A or B) *best* shows: **Marge expects to be reading a book.**



A B

Question 3:

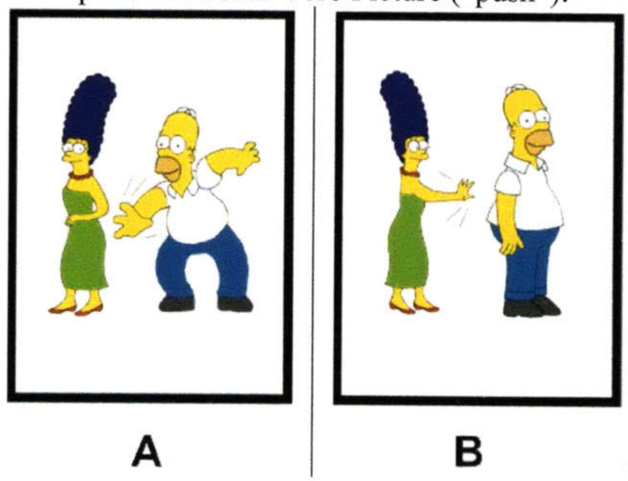
Which picture (A or B) *best* shows: **Lisa expects to be washing the floor.**



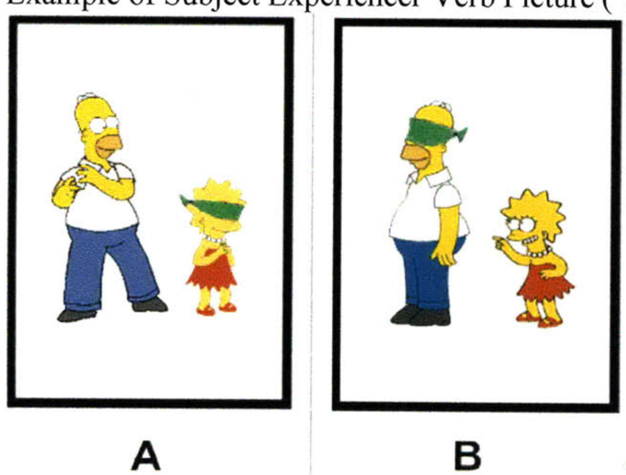
A B

A4: Study 3

Example of Actional Verb Picture ("push"):



Example of Subject Experiencer Verb Picture ("see"):



Test Items (as Linger file).

Guide to test items (as Linger file)
Line 1: # <study name> <item number> <condition>
Line 2: <image file>
Line 3: <test sentence>
Line 4: <side of screen on which correct choice appears>

Conditions:
AA: Actional Active
ALP: Actional Long Passive (i.e. with by-phrase)
ASP: Actional Short Passive (i.e. no by-phrase)
PA: Psychological Active (i.e. Subject Experiencer Active)
PLP: Psychological Long Passive (i.e. SE Passive with by-phrase)

PSP: Psychological Short Passive (i.e. SE Passive with no by-phrase)

Test items:

Study-3 1 AA
Images/1.jpg
Homer pushes Marge.
L

Study-3 2 AA
Images/2.jpg
Lisa pushes Bart.
R

Study-3 3 AA
Images/3.jpg
Homer kicks Lisa.
L

Study-3 4 AA
Images/4.jpg
Marge kicks Bart.
R

Study-3 5 AA
Images/5.jpg
Lisa kisses Homer.
L

Study-3 6 AA
Images/6.jpg
Bart kisses Marge.
R

Study-3 7 AA
Images/7.jpg
Marge holds Homer.
L

Study-3 8 AA
Images/8.jpg
Bart holds Lisa.
R

Study-3 9 ALP
Images/9.jpg

Bart is pushed by Lisa.
L

Study-3 10 ALP
Images/10.jpg
Marge is pushed by Homer.
R

Study-3 11 ALP
Images/11.jpg
Bart is kicked by Marge.
L

Study-3 12 ALP
Images/12.jpg
Lisa is kicked by Homer.
R

Study-3 13 ALP
Images/13.jpg
Marge is kissed by Bart.
L

Study-3 14 ALP
Images/14.jpg
Homer is kissed by Lisa.
R

Study-3 15 ALP
Images/15.jpg
Lisa is held by Bart.
L

Study-3 16 ALP
Images/16.jpg
Homer is held by Marge.
R

Study-3 17 ASP
Images/17.jpg
Lisa is pushed.
L

Study-3 18 ASP
Images/18.jpg
Homer is pushed.

R

Study-3 19 ASP
Images/19.jpg
Marge is kicked.
L

Study-3 20 ASP
Images/20.jpg
Homer is kicked.
R

Study-3 21 ASP
Images/21.jpg
Bart is kissed.
L

Study-3 22 ASP
Images/22.jpg
Lisa is kissed.
R

Study-3 23 ASP
Images/23.jpg
Bart is held.
L

Study-3 24 ASP
Images/24.jpg
Marge is held.
R

Study-3 25 PA
Images/25.jpg
Lisa remembers Homer.
L

Study-3 26 PA
Images/26.jpg
Bart remembers Marge.
R

Study-3 27 PA
Images/27.jpg
Marge hates Homer.
L

Study-3 28 PA
Images/28.jpg
Bart hates Lisa.
R

Study-3 29 PA
Images/29.jpg
Homer loves Marge.
L

Study-3 30 PA
Images/30.jpg
Lisa loves Bart.
R

Study-3 31 PA
Images/31.jpg
Homer sees Lisa.
L

Study-3 32 PA
Images/32.jpg
Marge sees Bart.
R

Study-3 33 PLP
Images/33.jpg
Marge is remembered by Bart.
L

Study-3 34 PLP
Images/34.jpg
Homer is remembered by Lisa.
R

Study-3 35 PLP
Images/35.jpg
Lisa is hated by Bart.
L

Study-3 36 PLP
Images/36.jpg
Homer is hated by Marge.
R

Study-3 37 PLP
Images/37.jpg
Bart is loved by Lisa.
L

Study-3 38 PLP
Images/38.jpg
Marge is loved by Homer.
R

Study-3 39 PLP
Images/39.jpg
Bart is seen by Marge.
L

Study-3 40 PLP
Images/40.jpg
Lisa is seen by Homer.
R

Study-3 41 PSP
Images/41.jpg
Bart is remembered.
L

Study-3 42 PSP
Images/42.jpg
Lisa is remembered.
R

Study-3 43 PSP
Images/43.jpg
Bart is hated.
L

Study-3 44 PSP
Images/44.jpg
Marge is hated.
R

Study-3 45 PSP
Images/45.jpg
Lisa is loved.
L

Study-3 46 PSP

Images/46.jpg
Homer is loved.
R

Study-3 47 PSP
Images/47.jpg
Marge is seen.
L

Study-3 48 PSP
Images/48.jpg
Homer is seen.
R

A5: Study 4

Notes: Sample test scenes (pictures) shown in Section 3.2.2

Test items (actual testing sheet):

Study 4

Kid's Name: Day Care: DOB: DT: Age:

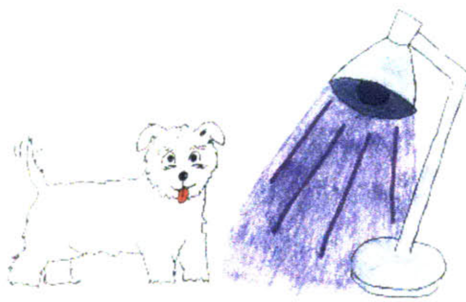
#	Scenario	Sentence	Ans	T/F
0	Mr. Bear calls tester incorrect sex	Jeff is a girl. (Robyn is a boy)		F
0	Barbie calls tester incorrect sex	Barbie thinks that Jeff is a girl (Robyn is a boy)		T
0	Barbie is wearing her scarf, but doesn't know it	Barbie thinks that she's wearing a scarf.		F
1	Ken puts pineapple in Barbie's bag. She doesn't know.	Barbie thinks that she is carrying a pineapple.		F
2	Barbie eats apple. Thinks orange. Ken knows apple	It seems to Ken that Barbie is eating an apple		T
3	Barbie's in Bedroom. Thinks outside. Ken thinks kitchen.	It seems to Ken that Barbie is in the bedroom.		F
4	Barbie has a necklace. Doesn't know. Ken does	Ken seems to Barbie to be wearing a necklace		F
5	Barbie is at school. Thinks she's home. Ken knows she's at school.	Ken thinks that Barbie is at school.		T
6	Ken puts pencil in B's bag, but thinks pen. She doesn't know	It seems to Ken that Barbie is carrying a pencil		F
7	Barbie has belt. Doesn't know. Ken does.	Barbie seems to be wearing a belt.		T
8	Barbie has a red backpack. Thinks blue. Ken knows red	Barbie seems to Ken to be carrying a red backpack		T
9	Barbie has glasses. Doesn't know. Ken does.	Barbie thinks that she is wearing glasses.		F
10	Ken doesn't put a book in Barbie's bag. She thinks she has one	Barbie seems to be carrying a book		F
11	Barbie is in woods. Thinks in city. Ken knows in woods.	Barbie seems to Ken to be in the woods.		T
12	Barbie has hat. Doesn't know. Ken thinks is scarf	It seems to Ken that Barbie is wearing a hat		F
13	Ken puts pineapple in Barbie's bag. She doesn't know.	Ken thinks that Barbie is carrying a pineapple.		T
14	Barbie doesn't have belt. Thinks she does. Ken knows she doesn't	Barbie seems to be wearing a belt		F
15	Ken puts pencil in Barbie's bag. She doesn't know.	It seems to Ken that Barbie is carrying a pencil.		T

16	Barbie is in mountains. Thinks at beach. Ken knows in mts.	Barbie seems to be at the beach.		F
17	Barbie has a hat. Doesn't know. Ken does.	It seems to Ken that Barbie is wearing a hat.		T
18	Ken puts candy in Barbie's bag. She doesn't know.	Barbie seems to Ken to be carrying some candy.		T
19	Barbie riding cow. Thinks is horse. Ken knows is cow.	Barbie seems to be riding a cow		T
20	Barbie is at school. Thinks she's home. Ken knows she's at school.	Barbie thinks that she is at school.		F
21	Ken puts book in Barbie's bag. She doesn't know.	Barbie seems to be carrying a book.		T
22	Barbie has a red backpack. Thinks blue. Ken knows red.	Ken seems to Barbie to be carrying a red backpack.		F
23	Barbie has glasses. Doesn't know. Ken does.	Ken thinks that Barbie is wearing glasses.		T
24	Barbie is in woods. Thinks in city. Ken knows in woods.	Ken seems to Barbie to be in the woods.		F
25	Barbie's in Bedroom. Thinks outside. Ken knows bedroom.	It seems to Ken that Barbie is in the bedroom.		T
26	Barbie has a necklace. Doesn't know. Ken does.	Barbie seems to Ken to be wearing a necklace.		T
27	Barbie eats apple. Thinks orange. Ken thinks is banana.	It seems to Ken that Barbie is eating an apple.		F
28	Barbie is at beach. Thinks in mountains. Ken knows at beach.	Barbie seems to be at the beach.		T
29	Barbie has orange umbrella. Thinks blue. Ken knows orange.	Ken thinks that Barbie is holding a green umbrella		T
30	Ken puts candy in Barbie's bag. She doesn't know.	Ken seems to Barbie to be carrying some candy.		F
31	Barbie riding horse. Thinks is cow. Ken knows is horse.	Barbie seems to be riding a cow		F
32	Barbie has orange umbrella. Thinks blue. Ken knows orange.	Barbie thinks that she is carrying a green umbrella		F

A6: Study 5

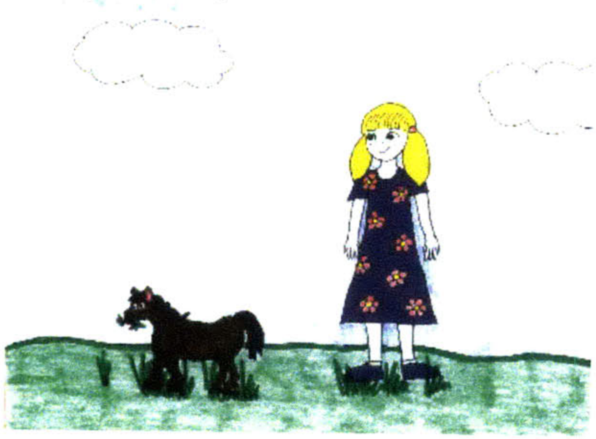
Picture Sets:

Reality ("white")



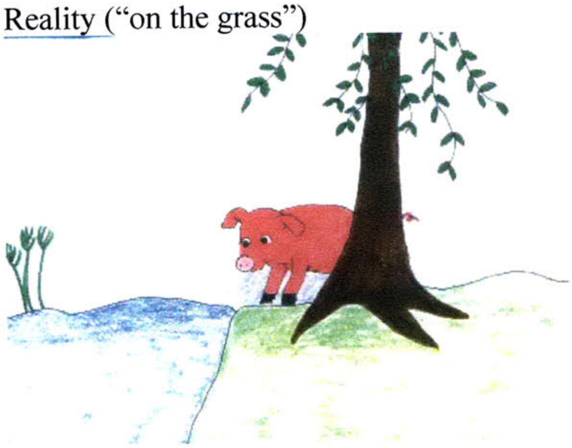
Appearance ("purple")

Reality ("small")

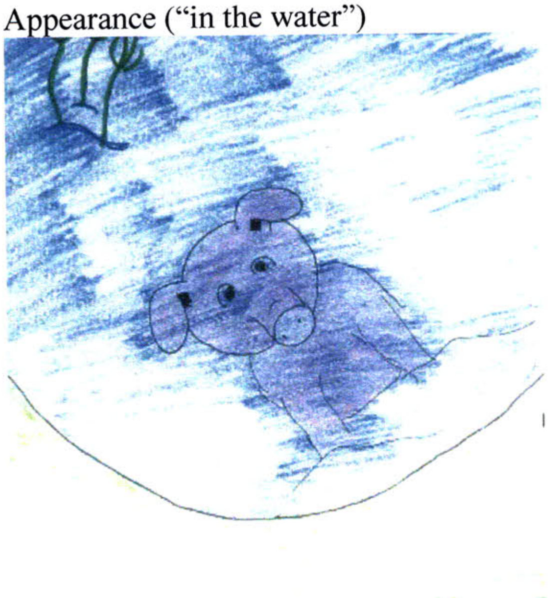


Appearance ("big") [requires use of magnifying glass]

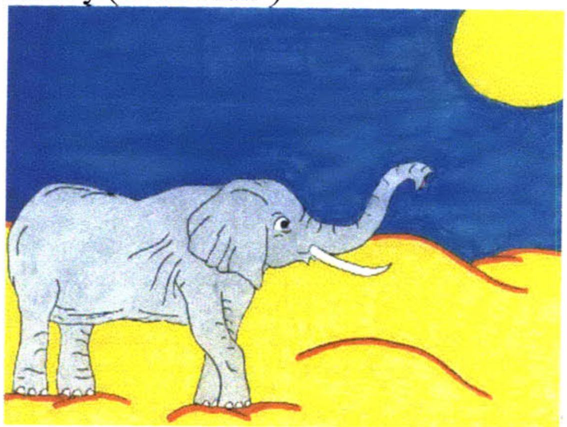
Reality ("on the grass")



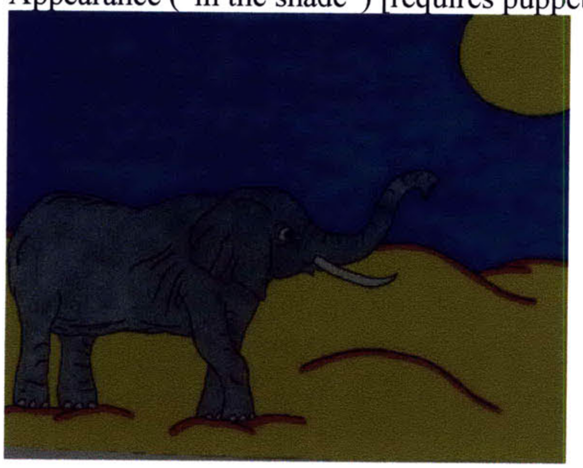
Appearance ("in the water")



Reality ("in the sun")



Appearance ("in the shade") [requires puppet to wear sunglasses]



Test Items (actual testing sheet):

Study 5

Kid's Name: Day Care: DOB: DT: Age:

#	Scenario	Sentence	Answer	T/F
0	Pre-test	The dog is purple		0
0	Pre-test	The dog looks purple		1
1	Horse	The horse really is big		0
2	Elephant	It really seems that the elephant is in the sun		0
3	Pig	The pig really likes to be in the water		0
4	Elephant	The elephant really seems to be in the shade		1
5	Dog	The dog really likes to be white		1
6	Horse	The horse really hates to be small		0
7	Dog	It really seems that the dog is purple		1
8	Pig	The pig really is on the grass		1
9	Pig	The pig really appears to be on the grass		0
10	Dog	The dog really is purple		0
11	Elephant	The elephant really hates to be in the shade		1
12	Horse	The horse really appears to be small		0
13	Pig	It really appears that the pig is on the grass		0
14	Elephant	The elephant really is in the shade		0
15	Horse	It really appears that the horse is big		1
16	Dog	The dog really seems to be purple		1

17	Pig	The pig really is in the water		0
18	Elephant	It really seems that the elephant is in the shade		1
19	Pig	The pig really likes to be on the grass		1
20	Dog	The dog really seems to be white		0
21	Elephant	The elephant really is in the sun		1
22	Horse	The horse really hates to be big		1
23	Horse	It really appears that the horse is small		0
24	Dog	The dog really likes to be purple		0
25	Dog	The dog really is white		1
26	Horse	The horse really appears to be big		1
27	Horse	The horse really is small		1
28	Pig	The pig really appears to be in the water		1
29	Elephant	The elephant really seems to be in the sun		0
30	Dog	It really seems that the dog is white		0
31	Pig	It really appears that the pig is in the water		1
32	Elephant	The elephant really hates to be in the sun		0

References

Anagnostopoulou, E. 1997. Clitic Left Dislocation and Contrastive Left Dislocation. In E. Anagnostopoulou, H. van Riemsdijk and F. Zwarts (eds.) *Materials on Left Dislocation*. Amsterdam: John Benjamin. 151-192.

Anagnostopoulou, E. 2003. Particles and voice. In A. Alexiadou, M. Rathert and A. von Stechow (eds.) *Perfect Explorations*. Berlin/New York: Mouton de Gruyter.

Babyonyshev, M., and D. Brun. 2004. The acquisition of perfective and imperfective passive constructions in Russian. *University of Pennsylvania Working Papers in Linguistics* 10, 1, 17–31.

Babyonyshev, M., J. Ganger, D. Pesetsky, and K. Wexler. 2001. The maturation of grammatical principles: Evidence from Russian unaccusatives. *Linguistic Inquiry* 32, 1, 1-44.

Baker, C., and J. McCarthy. 1981. *The logical problem of language acquisition*. Cambridge, MA: MIT Press.

Baker, M. 1988. *Incorporation: A theory of grammatical function changing*. Chicago: The University of Chicago Press.

Baltin, M. 2000. A-Movement. In M. Baltin and C. Collins (eds.) *The Handbook of Contemporary Syntactic Theory*. Oxford: Blackwell Publishers.

Barss, A., and H. Lasnik. 1986. A note on anaphora and double objects. *Linguistic Inquiry* 17, 219-251.

Bartke, S. 2004. Passives in German children with WS. In S. Bartke and Siegmuller (eds.) *Williams syndrome across languages*. Amsterdam/Philadelphia.

Becker, M. 2004. The search for structure: How learners distinguish raising from control. Handout from talk given to the Linguistics Department at the University of Maryland. (5/2004).

Becker, M. 2005. Raising, control and the subset principle, to appear in *Proceedings of WCCFL 24*, Somerville, MA: Cascadilla Proceedings Project, 52–60.

Becker, M. 2006. There began to be a learnability puzzle, to appear in *Linguistic Inquiry* 37, 3, 441-456.

Belletti, A., and L. Rizzi. 1981. The syntax of *ne*: some theoretical implications. *Linguistic Review* 1, 117–154.

Belletti, A., and L. Rizzi. 1988. Psych-verbs and θ-theory. *Natural Language and Linguistic Theory* 6, 291-352.

Berman, R. 1985. Acquisition of Hebrew. In D. Slobin (ed.) *The crosslinguistic study of language acquisition, Vol I*. Hillsdale, NJ: Erlbaum. 255-371.

Bever, T. 1970. The cognitive basis for linguistic structures. In Hayes, J. (ed.) *Cognition and the Development of Language*, New York: Wiley. 279-362.

Bishop, D., C. Adams, and C. Norbury. 2006. Distinct genetic influences on grammar and phonological short-term memory deficits: evidence from 6-year-old twins. *Genes, Brain, and Behavior*. Early online article.

Boeckx, C. 1998. Agreement constraints in Icelandic and elsewhere. *Working Papers in Scandinavian Syntax* 62, 1-35.

Boeckx, C. 1999. Conflicting C-command Requirements. *Studia Linguistica* 53, 227-250.

Boeckx, C., and K. Grohmann. 2007. Remark: Putting phases in perspective. *Syntax* 10:2, 204-222.

Boeckx., C., and N. Hornstein. 2003. Reply to control is not movement. *Linguistic Inquiry* 34, 269-280.
Borer, H., and K. Wexler. 1987. The maturation of syntax. In T. Roeper and E. Williams (eds.) *Parameter Setting*. Dordrecht: Reidel.
Borer, H. and K. Wexler. 1992. Bi-unique relations and the maturation of grammatical principles. *Natural Language and Linguistic Theory* 10, 147-189.
Borsley, R. 1999. *Syntactic theory: A unified approach.* London and New York: Arnold.
Boysson-Bardies, B. 1999. *How language comes to children.* Cambridge, MA: MIT Press.
Bresnan, J. 1982. The passive in lexical theory. In J. Bresnan (ed.) *The mental representation of grammatical relations*. Cambridge, MA: MIT Press. 3–86.
Cameron-Faulkner, T., E. Lieven, and M. Tomasello. 2003. A construction based analysis of child directed speech. *Cognitive Science* 27, 843-873.
Chang, H. 1986. Young children's comprehension of the Chinese passives. In H. Kao and R. Hoosain (eds.) *Linguistics, Psychology and the Chinese Language*. Hong Kong: Center of Asian Studies, University of Hong Kong.
Chien, Y.-C. & K. Wexler. 1990. Children's knowledge of locality conditions in binding as evidence for the modularity of syntax and pragmatics. *Language Acquisition* 1, 225-295.
Chomsky, C. 1969. *The acquisition of syntax in children from 5 to 10.* Cambridge: MIT Press.
Chomsky, N. 1959. Review of Verbal Behavior by B. F. Skinner. *Language*, 35, 26-58.
Chomsky, N. 1981. *Lectures on government and biding.* Dordrecht: Foris.
Chomsky, N. 1986. *Knowledge of language: Its nature, origin, and use.* New York: Praeger.
Chomsky, N. 1995. *The Minimalist Program.* Cambridge, MA: MIT Press.
Chomsky, N. 1998. Minimalist inquiries: The framework. In R. Martin, D. Michaels, and J. Uriagerkda (eds.) *Step by step: Essays on Minimalist syntax in honor of Howard Lasnik.* Cambridge, MA: MIT Press.
Chomsky, N. 2001. Derivation by Phase. In M. Kenstowicz (ed.) *Ken Hale: A Life in language*. Cambridge, MA: MIT Press. 1-52.
Collins, C. 2005a. A Smuggling approach to raising in English. *Linguistic Inquiry* 36, 2, 289-298.
Collins, C. 2005b. A Smuggling approach to the passive in English. *Syntax* 8, 2, 81-120
Collins, C., and H. Thráinsson. 1996. VP-internal structure and object shift in Icelandic. *Linguistic Inquiry* 27, 33, 391-444.
Crain, S., and J. Fodor. 1993. Competence and performance in child language in E. Dromi (ed.) *Language and cognition: A developmental perspective.* Norwood: Ablex. 141-171.
Crain, S., and M. Nakayama. 1987. Structure dependence in grammar formation. *Language*. 63, 522-543.
Crain, S., R. Thornton, and K. Murasugi 1987. Capturing the evasive passive. Paper presented at the 12th Annual Boston University Conference on Language Development.

Crain, S., and K. Wexler. 1999. Methodology in the study of language acquisition: A modular approach. In W. Ritchie and T. Bhatia (eds) *Handbook of Language Acquisition*. San Diego: Academic Press.

Crawford, J. 2005. The acquisition of the Sesotho passive: Reanalyzing a counterexample to maturation. ms. Boston University.

Crawford, J., and C. Hirsch. 2008. Acquisition of Sesotho Passives: Evidence for maturation. Paper presented at GALANA 3 University of Connecticut.

Culicover, P., and R. Jackendoff. 2001. Control is not movement. *Linguistic Inquiry* 30, 483-511.

Demuth, K. 1989. Maturation and the acquisition of the Sesotho passive. *Language* 65, 56-80.

Déprez, V., and A. Pierce. 1993. Negation and functional projections in early grammar. *Linguistic Inquiry* 24, 25-67.

Diderichsen, P. 2001. Selective attention in the development of the passive construction: a study of language acquisition in Danish children. In E. Engberg-Pedersen and P. Harder (eds.) *Ikonicitet og struktur. Netværk for Funktionel Lingvistik*, Department of English, University of Copenhagen.

Diesing, M. 1992. *Indefinites.* Cambridge, MA: MIT Press

Dikken, M. den. 2006. A reappraisal of vP being phasal – A reply to Legate. *Theoretical Linguistics*. Available from: http://web.gc.cuny.edu/dept/lingu/dendikken/docs/legate_reply.pdf

Djurkovic, M. 2005. Structural vs. frequency effects in L1 acquisition of the passive and impersonal in Serbian. Paper presented at the 27th Conference on Germanic Linguistics, Workshop on Frequency effects in language acquisition. Embick, D. 2003. Locality, listedness, and morphological identity. *Studia Linguistica* 57, 3, 143–170.

Embick, D. 2004. On the structure of resultative participles in English. *Linguistic Inquiry* 35, 3, 355-392.

Epstein, S., E. Groat, R. Kawashima, and H. Kitahara, 1998. *A derivational approach to syntactic relations.* Oxford: Oxford University Press.

Ferguson, K. 1994. Deriving the invisibility of PP-nodes for Command from AGR-P+P Case checking. In S. Epstein, H. Thrainsson and S. Kuno (eds.) *Harvard working papers in linguistics* 4. Cambridge, MA: Harvard University Press.

Ferguson, K. 1996. A feature-relativized shortest move requirement. Doctoral dissertation. Harvard University.

Fodor, J. 1966. How to learn to talk: Some simple ways. In F. Smith and G. Miller (eds.) *The genesis of language.* Cambridge, MA: MIT Press.

Fox, D., and Y. Grodzinsky. 1998. Children's passive: A view from the *by*-phrase. *Linguistic Inquiry* 29, 311-332.

Froud, K., K. Wexler, and V. Tsakali. In preparation. The development of raising. ms. MIT.

Gabriel, R. 2001. A aquisição das construções passivas em Português e Inglês: Um estudo translinguistico. Doctoral dissertation. PUCRS, Porto Alegre (Brazil).

Ganger, J., S. Dunn, and P. Gordon. 2004. Genes take over when the input fails: A twin study of the passive. Poster presented at BUCLD 28.

Gibson, E. 1998. Linguistic complexity: Locality of syntactic dependencies. *Cognition* 68, 1-76.

Gibson, E. 2000. The dependency locality theory: A distance-based theory of linguistic complexity. In Y. Miyashita, A. Marantz, and W. O'Neil (eds.) *Image, language, brain.* Cambridge, MA: MIT Press.

Goodluck, H., A. Terzi, and G. Chocano Díaz. 2001. The acquisition of control crosslinguistically: Structural and lexical factors in learning to license PRO. *Journal of Child Language* 28, 1, 153-172.

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

Gordon, P., R. Hendrick, and M. Johnson. 2001. Memory interference during language processing. *Journal of Experimental Psychology: Learning, Memory and Cognition* 27, 1411-1423.

Gordon, P., Hendrick, R, and W. Levine. 2002. Memory load interference during syntactic processing. *Psychological Science* 13, 425-420.

Grimm, H., H. Schöler, and M. Wintermantel. 1975. *Zur Entwicklung sprachlicher Strukturformen bei Kindern.* Weinheim: Beltz.

Guasti, M. 2002. *Language acquisition: The growth of grammar.* Cambridge, MA: MIT Press

Haegeman, L. 1994. *Introduction to government and binding.* Oxford: Blackwell.

Hamburger, H., and S. Crain. 1982. Relative acquisition. In S. Kuczaj (ed.) *Language Development Vol. I: Syntax and Semantics.* Hillsdale, NJ: Erlbaum.

Harley, H. 2007. External arguments: On the independence of Voice and *v*. Talk presented at GLOW XXX.

Harwood F. 1959. Quantitative study of the speech of Australian children. *Language and Speech* 2, 236–270.

Hicks, G. 2009. 'Tough'-constructions and their derivation. *Linguistic Inquiry*, 40, 4, 535-566.

Hirsch, C., and J. Hartman. 2006a. Some (*wh-*) questions concerning passive interactions. In A. Belletti, E. Bennati, C. Chesi, E. Di Domenico, and I. Ferrari (eds.) *Proceedings of the Conference on Generative Approaches to Language Acquisition.* Cambridge, UK: Cambridge Scholars Press.

Hirsch, C., and J. Hartman. 2006b. Semantic constraints on passive comprehension: Evidence for a theory of sequenced acquisition. Paper presented at Latsis Colloquium on Early Language Development and Disorders, Geneva, Switzerland.

Hirsch, C., N. Modyanova, and K. Wexler. 2006. Predicting passive acquisition: Environmental factors and biological maturation. Paper presented at Scandinavian Conference of Linguistics, Workshop in Language Acquisition.

Hirsch, C., R. Orfitelli and K. Wexler. 2007. When *seem* means *think*: The role of the experiencer-phrase in children's comprehension of raising. In A. Belikova, L. Meroni, and M. Umeda (eds.) *Proceedings of the 2nd Conference on Generative Approaches to Language Acquisition North America (GALANA).* 135-146.

Hirsch, C., R. Orfitelli, and K. Wexler. 2008. The acquisition of raising reconsidered. In A. Gavarro and M. Freitas (eds.) *Proceedings of the Conference on Generative Approaches to Language Acquisition.* Cambridge, UK: Cambridge Scholars Press.

Hirsch, C., and K. Wexler. 2006a. Children's passives and their *resulting* interpretation. In K. Deen, J. Nomura, B. Schulz, and B. Schwartz (eds.) *The Proceedings of the Inaugural Conference on Generative Approaches to Language Acquisition–North America, University of Connecticut Occasional Papers in Linguistics* 4. 125-136.

Hirsch, C., and K. Wexler. 2006b. By the way, children don't know *by*. In D. Bamman, T. Magnitskaia, and C. Zaller (eds.) *BUCLD 30: Proceedings of the 30th annual Boston University Conference on Language Development.* Somerville, MA: Cascadilla Press. 249-261.

Hirsch, C., and K. Wexler. 2006c. Acquiring verbal passives: Evidence for a maturational account. Paper presented at LSA Annual Conference.

Hirsch, C., and K. Wexler. 2007a. The late acquisition of raising: What children seem to *think* about *seem*. In S. Dubinsky and B. Davies (eds.) *New horizons in the analysis of control and raising*. New York: Springer.

Hirsch, C., and K. Wexler. 2007b. The development of inverse copulas (and clefts). Paper presented at Western Conference of Linguistics.

Hirsch, C. and K. Wexler. In preparation. Children's asymmetrical acquisition of clefts and relative clauses: Why processing and grammatical representation both matter.

Horgan, D. 1978. The development of the full passive. *Journal of Child Language* 5, 65-80.

Hornstein, N. 1999. Movement and control. *Linguistic Inquiry* 3, 69-96.

Hyams, N., D. Ntelitheos, and C. Manorohanta. 2006. The acquisition of the Malagasy voicing system. ms. UCLA.

Hyams, N., and W. Snyder. 2005. Young children never smuggle: Reflexive clitics and the universal freezing hypothesis. Paper presented at BUCLD.

Ito, M., and K. Wexler. 2002. The acquisition of Japanese unaccusatives. ms. MIT.

Johnson., K. 1991. Object positions. *Natural Language and Linguistic Theory* 9, 577-636.

Kayne, R. 1984. Principles of particle constructions. In J. Guéron, H-G. Obenauer, and J-Y. Polluck (eds.) *Grammatical representation.* Dordrecht: Foris. 101-140.

Kim, S. 2005. On the dative intervention effect on raising to subject. *Studies in Generative Grammar* 15, 4, 649-660.

Kirby, S. 2005. Semantics or subcases? The acquisition of referential vs. expletive *it.* Master thesis. Chapel Hill.

Kirby, S. 2009. Semantic scaffolding in first language acquisition: The acquisition of raising-to-object and object control. Doctoral dissertation. Chapel Hill.

Kirby, S. and M. Becker. 2007. Which *it* is it? The acquisition of referential and expletive *it*. *Journal of Child Language* 34, 571-599.

Kitihara, H. 1997. *Elementary operations and optimal derivations.* Cambridge, MA: MIT Press.

Koizumi, M. 1993. Object agreement phrases and the split VP hypothesis. In C. Phillips & J. Bobalijk (eds.) *Papers on Case and Agreement.* Cambridge, MA: MIT Press.

Koopman, H., and D. Sportiche. 1991. The position of subjects. *Lingua* 85, 211-258.

Koring, L., and K. Wexler. 2009. Children's Late Acquisition of Dutch Raising-Verbs. In Y. Otsu (ed.) *The Proceedings of the Tenth Tokyo Conference on Psycholinguistics.* Tokyo: Hituzi Syobo Publishing. 143-166.

Kratzer, A. 1996. Severing the external argument from its verb. In J. Rooryck and L. Zaring (eds.) *Phrase structure and the lexicon.* Dordrecht: Kluwer. 109–137.

Kratzer, A. 2000. Building statives. In L. J. Conathan (ed.) *Proceedings of the 26th Annual Meeting of the Berkeley Linguistics Society*. 385-399.

Landau, I. 2003. Movement out of control. *Linguistic Inquiry* 34, 471-498.

Larson, R. 1991. Promise and the theory of control. *Linguistic Inquiry* 22, 103–39.

Lasnik., H. 1995. Case and expletives revisited: on greed and other human failings. *Linguistic Inquiry* 26, 615-633.

Lasnik, H. 1998. On a scope reconstruction paradox. Celebration: An Electronic Festschrift in Honor of Noam Chomsky's 70th Birthday.

Lasnik, H., and M. Saito. 1991. On the subject of infinitives. In *Proceedings From the Annual Meeting of the Chicago Linguistic Society (CLS)* 27, 324-343.

Lee, H., and K. Wexler 2001. Nominative case omission and unaccusatives in Korean acquisition. In J. Ree (ed.) *The Proceedings of the International Conference on Korean Linguistics*. Prague. 263-279.

Legate, J. 2003. Phases in ''Beyond explanatory adequacy.'' ms. MIT.

Levin, B., and M. Rappaport. 1986. The formation of adjectival passives. *Linguistic Inquiry* 17, 623-661.

Levin, B., and M. Rappaport Hovav. 1995. *Unaccusivity: At the syntax-lexical semantics interface.* Cambridge, MA: MIT Press.

MacWhinney, B. 2000. *The CHILDES project: Tools for analyzing talk. Third edition.* Mahwah, NJ: Lawrence Erlbaum Associates.

Manzini, R., and K. Wexler. 1986. Parameters, binding theory and learnability. *Linguistic Inquiry* 18, 413-444.

Marantz, A. 1984. *On the nature of grammatical relations.* Cambridge, MA: MIT Press.

Maratsos, M., and R. Abramovitch. 1975. How children understand full, truncated, and anomalous passives. *Journal of Verbal Learning and Verbal Behavior* 14, 45-157.

Maratsos, M., D. Fox, J. Becker, and M. Chalkey. 1985. Semantic restrictions on children's passives. *Cognition* 19, 167-191.

Maratsos, M., S. Kuczaj, D. Fox, and M. Chalkley. 1979. Some empirical studies in the acquisition of transformational relations: Passives, negatives and the past tense. In W.A. Collins (ed.), *Children's Language and Communications.* Hillsdale, NJ: Lawrence Erlbaum Associates. 1-45.

Marchman, V., E. Bates, A. Burkardt, and A. Good. 1991. Functional constraints of the acquisition of the passive: toward a model of the competence to perform. *First Language* 11, 65-92.

McDaniel, D., J. Cairns, and J. Hsu. 1990. Binding principles in the grammars of young children. *Language Acquisition* 1, 121-138.

McGinnis, M. 1998. Locality in A-movement. Doctoral dissertation. MIT.

Menuzzi, S. 2001. The acquisition of passives in English and Brazilian Portuguese: regularities despite the input. In *Generative Approaches to Language Acquisition.* Lisbon: Associação Portuguesa de Lingüística. 172-179.

Miyagawa, S. 1989. *Structure and case-marking in Japanese.* New York: Academic Press.

Miyamoto, E., K. Wexler, T. Aikawa, and S. Miyagawa 1999. Case-dropping and unaccusatives in Japanese acquisition. In *Proceedings of the 23rd Annual Boston University Conference on Language Development.* Somerville, MA: Cascadilla Press. 443-452.

Müller, G. 1998. *Incomplete category fronting*. Dordrecht: Kluwer.
Orfitelli, R., and N. Hyams. 2008. An Experimental Study of Children's Comprehension of Null Subjects: Implications for Grammatical/Performance Accounts. In H. Chan, H. Jacob, and E. Kapia (eds.) *Proceedings of the 32nd annual Boston University Conference on Language Development*. 335-346.
Otsu, Y. 1994 Early acquisition of scrambling in Japanese. In T. Hoekstra and B. Schwartz (eds.) *Language Acquisition Studies in Generative Grammar*. John Benjamins. 253-264.
Perovic, A., and K. Wexler. 2007. Knowledge of binding, raising and passives in Williams Syndrome. In A. Belikova, L. Meroni, and M. Umeda (eds.) *Proceedings of the 2nd Conference on Generative Approaches to Language Acquisition North America (GALANA)*. 273-284
Pesetsky, D. 1982. Paths and categories. Doctoral dissertation. MIT.
Pesetsky, D. 1995. *Zero syntax: Experiencers and cascades*. Cambridge, MA: MIT Press.
Phillips, C. 1996. Order and structure. Doctoral dissertation. MIT.
Pierce, A. 1992a. *Language acquisition and syntactic theory: A comparative analysis of French and English child grammar*. Dordrecht: Kluwer.
Pierce, A. 1992b. The acquisition of passive in Spanish and the question of A-chain maturation. *Language acquisition* 2, 55-81.
Pinker, S. 1984. *Language learnability and language development*. Cambridge, MA: Harvard University Press.
Poeppel, D., and K. Wexler. 1993. The full competence hypothesis of clause structure in early German. *Language* 69, 1-33.
Postal, P. 1974. *On raising*. Cambridge, MA: MIT Press.
Pye, C., and P. Poz. 1988. Precocious Passives (and antipassives) in Quiche Mayan. *Papers and Reports on Child Language Development* 27:71-80.
Pye, C. 1992. The Acquisition of K'iche' Maya. In D. Slobin (ed.) *The crosslinguistic study of language acquisition*, Vol. 3. Hillsdale, NJ: Lawrence Erlbaum.
Reinhart, T., and T. Siloni. 2004. Against the unaccusative analysis of reflexives. In A. Alexiadou, E. Anagnostopoulou, and M. Everaert (eds.) *The unaccusativity puzzle: Explorations of the syntax–lexicon interface*. Oxford: Oxford University Press. 159-180.
Rizzi, L. 1990. *Relativized minimality*. Cambridge, MA: MIT Press.
Rizzi, L. 2001 Relativized Minimality effects. In M. Baltin and C. Collins (eds.) *The handbook of contemporary syntactic theory*. Oxford: Blackwell. 89-110.
Rizzi L., and U. Shlonsky. 2006. Satisfying the subject criterion by a non subject: English locative inversion and heavy NP subject shift. In M. Frascarelli (ed.) *Phases of interpretation*. Berlin: Mouton de Gruyter. 341-361.
Rosenbaum, P. 1967. *The grammar of English predicate complement constructions*. Cambridge, MA: MIT Press.
Runner, J. 1995. Noun phrase licensing and interpretation. Doctoral dissertation. University of Massachusetts, Amherst.
Runner, J. 2006. Lingering challenges to the raising to object and object control constructions. *Syntax* 9:2, 193-213.
Santelmann, L. 1995. *The acquisition of verb second grammar in child Swedish*. Doctoral dissertation. Cornell University.

Sano, T. 2000. Issues on unaccusatives and passives in the acquisition of Japanese. In *Proceedings of the Tokyo Conference on Psycholinguistics.* 1-21.
Schäfer, F. 2008. Two types of external argument licensing: The case of causers. Talk presented at GLOW XXXI.
Schütze, C., and K. Wexler. 1996. Subject case licensing and English root infinites. *Proceedings of the 10th Annual Boston University Conference on Language Development.* Somerville, MA: Cascadilla Press. 670-681.
Sherman, J. 1983. The acquisition of control in complement sentences: The role of structural and lexical factors. Doctoral dissertation. Cornell University.
Sinclair, H., and J. Bronckart. 1972. SVO—a linguistic universal. *Journal of Experimental Child Psychology* 14, 329-348.
Slobin, D. 1966. Grammatical transformations and sentence comprehension in childhood and adulthood. *Journal of Verbal Learning and Verbal Behavior* 5, 219-227.
Snyder, W., N. Hyams, and P. Crisma. 1995. Romance auxiliary selection with reflexive clitics: evidence for early knowledge of unaccusativity. In E. Clark (ed.) *Proceedings of the 26th annual Child Language Research Forum.* Stanford, CA: CSLI Publishings.
Stojanovic, D. 2004. Sources of difficulty with passives in child Serbian: A-movement or argument saturation? Poster presented at Formal approaches to Slavic linguistics (FASL) 13.
Stromswold, K. 1995. The acquisition of subject and object wh-questions. *Language Acquisition* 4, 5-48.
Stromswold, K. 1996. Does the VP-internal subject stage really exist? Paper presented at the 21st Annual Boston University Conference on Language Development.
Sudhalter, V., and M. Braine. How does comprehension of passive develop? A comparison of actional and experiential verbs. *Journal of Child Language* 12, 455-470
Sugisaki, K. 1998. Japanese passives in acquisition. In *University of Connecticut Working Papers in Linguistics* 10.
Tada, H. 1993. *A/A'* partition in derivation. Doctoral dissertation. MIT.
Terzi, A., and K. Wexler. 2002. A-chains and s-homophones in children's grammar: Evidence from Greek passives. In *NELS 32.*
Thráinsson, H. 1979. *On complementation in Icelandic.* New York: Garland.
Torrego, E. 1996. Experiencers and raising verbs. In Robert Freidin (ed.) *Current issues in comparative grammar.* Dordrecht: Kluwer.
Torrego, E. 2002. A derivational approach to syntactic relations based on clitics. In S. Epstein and T. Seely (eds.) *Derivation and explanation in the minimalist program.* Oxford: Blackwell.
Turner, E., and R. Rommetveit. 1967. Experimental manipulation of the production of active and passive voice in children. *Language and Speech* 10, 169-180.
Verrips, M. 1996. Potatoes must peel: The acquisition of the Dutch passive. Doctoral dissertation. University of Amsterdam.
Warren, T., and E. Gibson. 2002. The influence of referential processing on sentence complexity. *Cognition* 85, 79-112.
Wasow, T. 1977. Transformations and the Lexicon. In P. Culicover, T. Wasow and A. Akmajian (eds.) *Formal Syntax.* New York: Academic Press.

Watt, W. 1970. On two hypotheses concerning psycholinguistics. In J.R. Hayes (ed.) *Cognition and the development of language.* New York: Wiley.

Wellman, H., D. Cross, and J. Watson. 2001. Meta-analysis of theory-of-mind development: The truth about false belief. *Child Development* 72, 655-684.

Wellman, H., M. Hollander, and C. Schult. 1996. Young children's understanding of bubbles and of thought. *Child Development* 67, 768-788.

Wells, G. 1979. Variation in child language. In P. Fletcher and M. Garman (eds.) *Language acquisition.* New York: Cambridge Unsiversity Press.

Wexler, K. 1992. Some issues in the growth of control. In R.K. Larson, S. Iatridou, I. Lahiri, and J. Higginbotham (eds.) *Control and Grammar.* Dordrecht: Kluwer.

Wexler, K. 1993. Optional infinitives, head movement, and the economy of derivations. In D. Lightfoot, and N. Hornstein (eds.) *Verb Movement.* Cambridge: Cambridge University Press.

Wexler, K. 1998. Very early parameter setting and the unique checking constraint: A new explanation of the optional infinitive stage. *Lingua* 106, 23-79.

Wexler, K. 2002. Lenneberg's dream: Learning, normal language development and specific language impairment. In J. Schaeffer & Y. Levy (eds.) *Language Competence across Populations: Towards a definition of Specific Language Impairment.* Mahwah, NJ: Erlbaum. 11-61.

Wexler, K. 2004. Theory of phasal development: Perfection in child grammar. In *MIT Working Papers in Linguistics* 48, 159-209.

Wexler, K. and Y-C. Chien. 1985. The development of lexical anaphors and pronouns. *Papers and Report on Child Language Development* 24, 138-149.

Wexler, K. and P. Culicover. 1980. *Formal Principles of Language Acquisition.* Cambridge, MA: MIT Press.

Fin