LEARNING LOCATIVE VERBS:
HOW UNIVERSAL LINKING RULES CONSTRAIN PRODUCTIVITY

by

JESS HARRY GROPEN

B.A., Pomona College
(1982)

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

DOCTOR OF PHILOSOPHY

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

February 1989

Signature of Author____
Department of Brain and Cognitive Sciences
January 10, 1989

Certified by_____
Steven Pinker
Associate Professor, Cognitive Science
Thesis Supervisor

Accepted by_____
Emilio Bizzi
Head, Department of Brain and Cognitive Sciences

LEARNING LOCATIVE VERBS:
HOW UNIVERSAL LINKING RULES CONSTRAIN PRODUCTIVITY

by

JESS HARRY GROPEN

Submitted to the Department of Brain and Cognitive Sciences on January 10, 1989
in partial fulfillment of the requirements for the Degree of
Doctor of Philosophy in Cognitive Science

ABSTRACT

Bowerman (1982) documents a pattern of syntactic errors that children make in learning locative verbs--verbs like *pour* and *fill*, which express the transfer of some content to or from some container. According to Bowerman, children between the ages of four and seven often over-regularize the locative form with the content encoded as direct object, producing sentences such as **I filled the water into the glass* (cf. *I filled the glass with water*). Why does this pattern of errors occur, and how does the child ultimately *un*learn these errors? An account of locative verb learning is proposed in which children learn the syntactic privileges of verbs on the basis of their meanings. In particular, it is hypothesized that children make use of a universal linking rule of Object Affectedness, according to which an argument is encodable as the direct object of a verb if the entity to which it corresponds is affected in the meaning of the verb. For example, the meaning of *fill* specifies the particular way in which the container is affected (i.e., it undergoes a change of state from being empty to being full), but does not specify the particular manner (e.g., pouring or dripping) in which the content is affected. The universal thus predicts that the container, but not the content, is encodable as the direct object of *fill*.

Six experiments test the hypothesis that children make use of the linking regularity, but that they must learn what counts as affected in the meanings of particular verbs. From this proposal it follows that children will be productive, producing forms they haven't heard in the input--and in fact, before they figure out what counts as affected they may overgenerate locative forms (i.e., produce ungrammatical sentences). Furthermore, it is predicted that instances of overgeneration should be associated with corresponding misinterpretations of verb meaning. In Experiments 1 and 2, the ability of children and adults to understand and produce common locative verbs was tested. In particular, sets of drawings were used to assess the subjects' understanding of fine-grained aspects of verb meaning. As predicted, it was found that children made syntactic errors, overgenerating the locatives of *fill* and *empty* with the content encoded as direct object (e.g., saying things like "I'm filling the water into the glass"); that children made semantic mistakes, misinterpreting the meanings of *fill* and *empty* as having something essential to do with the manner in which content changes location; and that instances of overgeneration were (weakly) associated with corresponding misinterpretations of verb

meaning (e.g., children who interpreted the action of filling as having something essential to do with the manner in which content changes location were likely to have uttered "I'm filling water into the glass"). These findings were interpreted as providing support for the hypothesis that verb meaning and syntax are linked in the lexicons of language learners, and it is concluded that misinterpretations of particular verbs, coupled with linking regularities, may account for the occurrence of the syntactic errors.

In Experiments 3-6, the correspondence between verb syntax and semantics was tested in a more direct fashion. Children and adults were taught novel verbs in a neutral syntax (e.g., *this is mooping*), and then tested on their willingness to encode the content or the container as the direct object of the verb. In these experiments, the semantics of the novels verbs was an independent variable: the meanings of the verbs varied according to whether the content or the container was affected in a particular and salient way (e.g., whether the content moved in a zig-zagging fashion, or whether the container changed color). It was predicted that children and adults should produce relatively more content locatives for verbs in which the content changes location in a particular manner, and relatively more container locatives for verbs in which the container changes state in a particular way. The results of each experiment confirmed this prediction, and were taken to support the conclusion that the Universal of Object Affectedness *must* be used under some circumstances.

In the General Discussion, the statement of the universal linking rule (given above) is defended and developed. A survey of the cross-linguistic literature supports the view that the affectedness of direct objects is a universal tendency, applying across domains in English and in other languages. Finally, two sources of mediation were argued to be relevant to the use of the linking rule: the operation of a *set* of linking rules in a domain and the clustering of the verbs of a domain into subclasses. Based on these two factors, a proposal is outlined of how a child may come to use a linking rule to predict the syntactic privileges of verbs.

Thesis Supervisor: Dr. Steven Pinker

Title: Associate Professor of Cognitive Science

Table of Contents

Acknowledgements

Ever since the first time that I met him, in my interview for graduate work at MIT, Steve Pinker has been a constant source of inspiration. He has taught me how to think about cognitive science through the abundance and clarity of his advice, and by the brilliance of his example. More than anyone else, he has changed the way I work, and the way I present my work to others.

Michelle Hollander and Richard Goldberg were simply the best research assistants that a graduate student could ask for. They contributed to the substance of this research in countless ways, by drawing out children in the performance of a task, by telling me when my experimental proposals or flights of theory were wrongheaded, and by just keeping me company.

Fellow graduate students Kyle Cave and Michael Tarr have been with me from the beginning, providing me with comradeship and answering every computer-related question that I could pose. Paul Bloom has touched me with his good humor and keen interest in theoretical issues. I've also benefitted from discussions with members of my dissertation committee--Kay Bock, Susan Carey, and Ken Wexler--even if they weren't always close at hand.

I would like to thank the directors, parents, and especially children of the following centers for their participation in this research: Angier After School Program, Bowen After School Care Program, Inc., Cambridge Nursery School, Central School, Children's Village, Inc., Creative Development Center, KLH Center, Leventhal-Sidman Jewish Community Center, MIT Summer Day Camp, Needham Children's Community Center, Newton Community Service Center, Newton-Wellesley Children's Corner, Plowshares Child Care and After School Programs, Recreation Place, Red Barn Nursery School, Rosary Academy Learning Center, Second Church Nursery School, Temple Beth Shalom Children's Center, Underwood After School Program, and Zervas After School Program.

For years I've wanted to thank Jay David Atlas, mentor and friend, who first sparked my interest in cognitive science. He guided me at a time when I needed it most.

To Gail, for her strength, intelligence, and love

To Loki, Suvi, Daphne, and Leia, for their sweet devotion

To my parents with love

General Introduction

The purpose of this work is to understand how children master the verbs in one domain. The domain we have chosen for study is that of locative verbs--such as *pour*, *fill*, *empty*, and *load*--verbs which express the transfer of some content to or from some container. In particular, we will propose an account of lexical learning in which verb syntax and semantics bear a certain correspondence to one another at various points in development. Although studies on the relation of syntax and semantic are as old as linguistics itself (Aristotle, *Metaphysics*), much recent work in linguistics (Carter (1976b), Ostler (1980), Rappaport & Levin (1986), Jackendoff (1987)) and psycholinguistics (Pinker, 1989) has posited the existence of universal linking regularities between semantic or thematic roles (e.g., agent, patient) and grammatical relations (subject, direct object). The present study provides much-needed experimental evidence on the existence of linking rules in the acquisition of English locative verbs. It is our belief that whatever issues are relevant to the acquisition of these verbs will be relevant to the acquisition of all verbs, across all languages.

We have chosen to study locative verbs for several reasons. Besides the fact that they are among the most common verbs in English, and they are learned early (Bowerman, 1982), we believe that they constitute a true domain of verbs--not merely because they may be studied as such, but because children learn them as such. By *domain*, we mean a class of verbs with shared semantic and syntactic properties. All locative verbs share the semantic property of expressing the transfer of content to a container (1a), as in *Mike filled the cup with water*, or the transfer of content from a container (1b), as in *Fred cleared the table of dishes*. Some locative verbs are inherently nondirectional (1c), as in *Betty poured water into/from the cup*, specifying either the goal container, the source container, or both.

(1a) Mike filled the cup with water.
Lloyd covered the spot with a towel.
George loaded the gun with ammo.
Dan stuffed the hamper with laundry.

(1b) Fred cleared the table of dishes.
Sally emptied the carton of ice cream.
Bob drained the sink of water.
Tess cleaned the draperies (of lint).

(1c) Betty poured water into/from the cup.
Gus dumped garbage into/from the can.
Tom dripped paint onto the floor/from the brush.
Sue spilled coffee on Ned/from her mug.

These verbs may be further subdivided as to whether the content (2a) or container (2b) must be encoded as direct object. We shall refer to these syntactic forms as *content* and *container* locatives, respectively. Some locative verbs (2c), which we shall call *alternators*, may accept either the content or the container as direct object. A comprehensive list of locative verbs, arranged according to syntactic privileges, is provided in the Appendix.

(2a) Betty poured water into the cup/*the cup with water.
Betty poured water from the cup/*the cup of water.
Gus dumped garbage into the can/*the can with garbage.
Gus dumped garbage from the can/*the can of garbage.
Tom dripped paint onto the floor/*the floor with paint.
Tom dripped paint from the brush/*the brush of paint.

Sue spilled coffee on Ned/*Ned with coffee.

Sue spilled coffee from her mug/*her mug of coffee.

(2b) Mike filled the cup with water/*water into the cup.

Lloyd covered the spot with a towel/*a towel over the spot.

(2c) George loaded the gun with ammo/ammo into the gun.

Dan stuffed the hamper with laundry/laundry into the hamper.

Fred cleared the table of dishes/dishes from the table.

Sally emptied the carton of ice cream/ice cream from the carton.

Bob drained the sink of water/water from the sink.

Tess cleaned the draperies (of lint)/lint from the draperies.

Several findings of Bowerman (1982) convinced us, early on, of the promise of studying this domain. On the basis of detailed diary studies of her own children, plus relevant data from six other children, Bowerman has documented a 'U'-shaped developmental pattern in the production of locatives: although children initially appear to be accurate with these verbs, errors emerge within the range of roughly four to seven years of age; after the age of seven, the errors decline. Bowerman found that the most frequent errors involve children overextending the content-locative form to verbs that ordinarily encode only the container as direct object, as in "I didn't fill water up to drink it" (Eva, 4;1). Less frequently, children demoted the container argument to oblique object, and the content argument was omitted. An example of the latter is "pinching on the balloon" instead of "pinching the balloon" (Christy, 4;2). More examples of both kinds of errors appear in Table 1. According to Bowerman, errors of the converse type--involving a *re*placed or *mis*placed content--are much rarer.

Table 1

Examples of Overgeneralization (Bowerman, 1982)

Errors with figure [f] as direct object and ground [g] as oblique object		
E (3;0)		I'm going to touch it [f] on your pants [g]
C (4;3)	M:	Simon says, "Touch your toes" [g]
	C:	To what? [interprets toes as f, is now looking for g] [Note: this is a comprehension error]
C (6;10)		Feel your hand [f] to that [g]
E (5;0)		Can I fill some salt [f] into the bear [g] [= a bear-shaped salt shaker]
E (4;5)		I'm going to cover a screen [f] over me [g]
C (4;9)		She's gonna pinch it [f] on my foot [g]
Errors with figure [f] as direct object and ground [g] omitted		
E (4;1)		I didn't fill water [f] up to drink it; I filled it [f] up for the flowers to drink it
E (4;11)		And I'll give you these eggs [f] you can fill up [giving M beads to put into cloth chicken-shaped container g]
E (5;3)		Terri said if this [= rhinestone on a shirt] were a diamond then people would be trying to rob the shirt [f]
Errors with ground [g] demoted to oblique object		
C (3;11)		Eva is just touching gently on the plant [g]
C (4;2)		Pinch on the balloon [g]
Errors with ground [g] as direct object		
E (2;11)		Mommy, I poured you [g]
E (4;11)		I don't want it [= toast] because I spilled it [g] of orange juice [f]

Note: figure [f] corresponds to content; ground [g] corresponds to container.

Bowerman's explanation for this pattern of development is that a process of reorganization, driven by the semantic interpretation of locative forms, is responsible for the errors. This process, she argues, is not unlike the familiar example of verb and noun inflectional morphology: in the case of locatives, children first use unanalyzed syntactic forms (*fill the glass*; cf. *broke*); then, after they discover the semantic correlates of the forms by abstracting the correspondence between grammatical and semantic relations on the basis of particular verbs, children overgeneralize the content locative (*fill the water*; cf. *breaked*); finally, the container-locative form returns. Bowerman suggests (following Talmy, 1972) that the content locative is overgeneralized more than the container locative because it is the dominant pattern in English for expressing locative events. Accordingly, the overregularization of container locatives is rare (e.g., "I spilled it [container] of orange juice"; Eva, 4;11) for the same reason that the overregularization of minor past tense inflections is rare (e.g., *brang* on the pattern of *sang*).

Instructive, also, are two explanations that Bowerman (1982) rejects. First, she rejects the explanation that children are just making speech errors--for example, spontaneously substituting *fill* for *pour* because they are related in meaning. As Bowerman argues, there are plenty of observed errors involving verbs with no obvious substitutes, and furthermore, some of these errors involve *comprehension*, not production (See Table 1). Second, and more importantly, she rules out a purely syntactic explanation for the overregularization of locatives, according to which the child would regard NP-V-NP_i-*with*-NP_j and NP-V-NP_j-*into/onto*-NP_i (NP-V-NP_i-*of*-NP_j and NP-V-NP_j-*from/out of*-NP_i) as interchangeable forms. Relevant here is that Bowerman did *not* find errors like *I read Mary with a book* (from *I read a book to Mary*) and *I ate a spoon into my pudding* (from *I ate my pudding with a spoon*), presumably because children know that *read* and *eat* don't take contents and containers as arguments: Mary (in **I*

read Mary with a book) doesn't count as a container; a spoon (in **I ate a spoon into my pudding*) doesn't count as content. Thus, children appear to be constrained in their overregularization of these forms by the semantic roles of the verb's arguments. (Bowerman doesn't speak of *contents* and *containers*, but of *Figures* and *Grounds*. We will use the more specific--and we think, more appropriate--terminology.) Locative verbs constitute a domain because children become sensitive to a domain of arguments--an argument space--which probably specifies the shape, size, and dimensionality of potential containers and the mass/count properties of potential contents. The word *become* is important here; in Bowerman's view, the lateness of the errors implies that the argument space of locative forms is not grasped from the beginning of language learning. (We shall elaborate on the notion of an argument space in the General Discussion.)

In general, Bowerman sees her (1982) work as a corrective against strict anti-whorfianism (or strict whorfianism, for that matter); she rejects the strongest assumption of "cognition-first" semantics in which "meaning in language...is isomorphic with the nonlinguistic way of viewing the world" (p.331). We agree with Bowerman that the lateness of the errors argues for the *necessity* of experience with language, and against the possibility that children map only pre-established meanings onto forms. In this paper, however, we take up the complementary position of arguing against the *sufficiency* of experience with language, and for the necessity of (linguistic) semantic universals. The postulation of linguistic universals is prompted by a limitation of Bowerman's study of locative verb errors: although Bowerman's account of reorganization explains how the errors arise, it cannot (and does not attempt to) explain how they eventually disappear. As we shall see, a plausible account of how the errors are unlearned will suggest a reconsideration of their source as well.

Consider the case of a child who, like Eva, utters a locative form of *fill* with the content encoded as direct object--for example, *I'm filling water into the cup*. Here, we may say that the content-locative form has been overregularized, just as the affixation of *-ed* onto verbs is overregularized to mark past tense. The analogy breaks down, however, when it comes to *un*learning the errors. The dis-analogy arises because whereas positive evidence provides feedback on the overregularization of obligatory rules like past tense inflection, it does not provide feedback on the overregularization of optional rules like locative formation. If a child overgeneralizes an obligatory rule by affixing the regular past tense marking *-ed* onto *break* to form ***breaked*, the child will receive positive evidence of the error; adult speech will provide an explicit contrast--*broke*--to the ungrammatical form. (The ability of the learner to recognize this exception, on the basis of positive evidence, also assumes some version of the Uniqueness Principle (Wexler and Culicover, 1981; Pinker, 1984). In this case, the learner must assume that the concept of BREAK + PAST TENSE is associated with one and only one form.) But if a child overgeneralizes an optional rule of locative formation, no amount of positive experience with the language will tell the child that (e.g.) *fill* only takes the container-locative form. This is because no context of language *demands* the utterance of the content locative in the same way that it demands the marking of past tense. For this reason, Baker (1979) considers the exceptions to optional rules to be "embarrassing" compared to the "benign" exceptions to obligatory rules.

Furthermore, the fact that a child never hears *fill* in the content locative cannot be taken by the child as ("indirect negative") evidence that the form is ungrammatical, upon pain of rejecting every unheard form as ungrammatical. Similarly, adults do not provide the child with *direct* negative evidence about which strings are not in the language (Brown and Hanlon, 1970; Hirsh-Pasek, Treiman, and Schneiderman, 1984; Pinker, 1989): parents rarely correct, or otherwise provide more subtle feedback concerning, the utterances of their children; when they

do, they are most often concerned with the truth value of the proposition expressed by the child's utterance, and only rarely with its ungrammaticality; and when they are concerned with ungrammaticality, children appear to be oblivious to the intended correction of form (Braine, 1971). More fundamentally, Pinker (1989) argues that even if negative evidence were available *and* useful *and* used, it seems unlikely to be *necessary* to language acquisition. The unavailability of negative as well as positive evidence about ungrammaticality, coupled with the productivity of locative formation (see Table 1) and the apparent arbitrariness of which verbs take which locative forms (e.g., *load* alternates, but *fill* does not), lead to an instance of learnability problem known as "Baker's Paradox": the child has no way of knowing which verbs are exceptional, and therefore no way of retreating from the false hypothesis of an overgeneral grammar to the true hypothesis of the correct grammar.

In our view, explanation in developmental psycholinguistics must be sensitive not only to the phenomena of child language, but also to the demands of learnability theory. Accordingly, Steven Pinker and his colleagues (Pinker, 1984; Pinker, Lebeaux, and Frost, 1987; Pinker, 1989; Gropen, Pinker, Hollander, Goldberg, and Wilson, 1989) have pursued a solution to this paradox as it pertains to partial generalizations in the English lexicon--especially, datives, passives, causatives, and locatives. The hypothesis of Constrained Productivity, in its most general form, states that children can--and to some extent must--learn the syntactic privileges of verbs on the basis of their meanings (or sounds). The hypothesis thus denies a critical assumption of Baker's Paradox that the syntactic privileges of verbs within a domain are arbitrary. As we have already seen, a coarse-grained semantic analysis of the argument space of locatives plausibly accounts for the *absence* of certain errors (e.g., **I read Mary with a book*). By contrast, the hypothesis under consideration states that a finer-grained semantic analysis ("within domain") accounts for the *unlearning* of certain errors (e.g., **I'm filling water into the cup*). The hypothesis also suggests to us that *mistakes* about the fine-grained

meanings of verbs might be the *source* of the syntactic overregularization; the learnability problem would be solved, on this account, as children revise their interpretations of verb meanings. In the remainder of the introduction, we shall first present a sub-hypothesis about the correspondence between verb meaning and syntax in universal grammar, then show how the universal applies to the adult lexicon (i.e., that the errors could be unlearned in principle), and finally make specific predictions about the occurrence of syntactic and semantic errors in child language.

Although the hypothesis of Constrained Productivity need not be centrally concerned with *universals* of language, recent versions of the hypothesis have posited that the child is born motivated looking for circumscribed correspondences between syntax and semantics. In the case of locatives, we shall entertain the following sub-hypothesis: that the capacity of the child to predict the syntactic privileges of a verb depends upon a universal linking rule of Object Affectedness, according to which

(3) an argument is encodable as the direct object of a verb if the entity to which it corresponds is affected in the meaning of the verb

This specific statement of the linking rule is motivated by the tendency, across languages, for affected entities to be encoded as direct objects (Moravcsik, 1978; Hopper & Thompson, 1980). In English locatives, the question is whether the linking rule of Object Affectedness determines word order--that is, whether a given verb will take the content locative, the container locative, or both. Also notice, in the statement of the linking rule, that affectedness is a *sufficient* condition on which arguments may be encodable as direct object by the verb, and furthermore, that the linking rule leaves open the possibility that more than one entity may be affected in the meaning of a verb (i.e., an alternator). In the General Discussion, we shall defend this statement of the rule and outline a proposal of how a child may come to use such a

linking rule to predict the syntactic privileges of verbs. For the present, we shall make the case for affectedness by example.

One important point before we proceed: if we think of the typical locative event, it might be argued (as does Bowerman, 1982) that both the substance and the location are to some extent affected; after all, the content changes location and the container changes state (e.g., from being empty to being full). The essential thing, however, is not what happens in the world, but rather *what the verb takes* to happen in the world. Roughly speaking, our story is that the child learns which verbs take which locative forms on the basis of the verb's meaning, which specifies--among other things--what essential changes of state or position an object must undergo in order for the verb to apply. In this regard, we must distinguish between three levels of description: syntactic (in *italics*); semantic; and cognitive, or "what happens in the world." The latter two levels of description have both been presented in plain text, for the reason that a particular description (e.g., the change in the state of a container from empty to full) is systematically ambiguous between intensional and extensional interpretations. Thus, in learning the meaning of a verb, a child must abstract from the events of a verb's usage, in which a container changes state, to the meaning of the verb, which specifies that change of state. Nonetheless, it is the description *at the semantic level* that is relevant to Object Affectedness. In addition, we make the distinction between components of meaning which are *essential* to the meaning of a verb (i.e., which are part of a partial decomposition of a verb's meaning), and those which are merely *typical* of contexts to which a verb applies.

Let's consider the verbs *fill*, *pour*, and *stuff*. Most adult speakers of English share the following intuitions: filling--essentially--tells you something about the change of state that a container undergoes; namely, from unfilled to full. It wouldn't be filling if the container ended up empty or, for that matter, 3/4-full. On the other hand, filling says nothing specific about the

change of location that a substance undergoes. One can certainly fill a glass by pouring water into it, but it would still be filling if the water dripped into the glass from a faucet. We can summarize these intuitions, then, by saying that the meaning of *fill* specifies the way in which the container is affected, but does not specify the way in which the content is affected. According to the universal, then, we should expect to find that the direct object of *fill* encodes only the container argument. And indeed it does: one can say *Sally filled the glass with water*, but not **Sally filled water into the glass*. Thus, the potential learnability problem of a child being unable to unlearn, and an adult saying, *I filled water into the glass* is averted.

This interpretation of filling is an instance of what has been called the *holistic interpretation*, according to which a container becomes totally involved in the change of state indicated by a locative verb if the argument to which it corresponds is encoded as the direct object of the verb (Anderson, S., 1971; Schwartz-Norman, 1976). Thus, *Bob loaded the cart with apples* implies that the capacity of the cart has been exhausted, but *Bob loaded the apples into the cart* does not. Indeed, we argue that the holistic interpretation is a special case of Object Affectedness. What's new here, however, is the application of Affectedness to *contents*. To take an example, the meaning of *pour* tells you something essential about the way a liquid moves through the air, in a cohesive stream. It wouldn't be pouring if one drop at a time changed location; that would be dribbling or dripping. Neither would it be pouring if an entire array of drops or particles changed location; that would be splashing or showering. On the other hand, pouring says nothing specific about a container or any other reference object. Certainly, one may pour water into a glass, but it would still be pouring if the water missed the glass entirely. We can summarize these intuitions, then, by saying that the meaning of *pour* specifies the way in which the content is affected, but does not specify the way in which the container is affected. According to the universal, then, we should expect to find that the direct object of *pour* encodes only the content argument: one can say *Sally poured water into the glass*, but not **Sally poured*

the glass with water. Notice that much the same explanation holds for the correspondence between meaning (1c) and syntax (2a) for the other inherently nondirectional locative verbs.

As a final example, the meaning of the verb *stuff* specifies both the particular change of location that the content undergoes and the particular change of state that the container undergoes. In fact, the manner and endstate of stuffing appear to be mutually constraining: in stuffing clothes into a hamper, for instance, the clothing must be forced into the hamper (perhaps to the extent that the clothing is compressed) *because* the capacity of the hamper is exhausted; conversely, the fact that the clothing must be forced into the hamper seems to imply that the hamper is already *stuffed* (perhaps to the extent that the hamper bulges). According to the universal, then, the direct object of *stuff* should encode either the content or the container. And this is what we find: *stuff* is an alternator.

As we suggested above, mistakes about the fine-grained meanings of verbs, such as the identification of which entity is affected, might be the source of the syntactic overregularization reported by Bowerman (1982). On the face of it, this proposal comports well with the observation made by Gentner (1978, 1982) that children are quite slow in fixing the standard (adult) meanings of verbs, compared with the meanings of nouns, and that children have more difficulty with functional components of verb meaning (e.g., changes of state) than perceptual/actional components of verb meaning (e.g., changes of location). For example, she found that children had more difficulty in learning the meaning of *mix*, which specifies that a homogenous combination of substances be the result of an action, than in learning the meanings of *shake* or *stir*, which specify the particular motions involved in an action. If we therefore assume that contents are identified as affected more easily than are containers (on perceptual/cognitive grounds), we have an explanation for *why* the content locative occurs, and is overregularized, more frequently than the container locative. Notice that, on this account, we

gain explanatory power at the expense of Bowerman's suggestion that children overgeneralize the content-locative form more than the container-locative form because the content locative occurs more frequently than the container locative. Even if this pattern exists, Bowerman is left with the question of *why* the content locative is dominant, or unmarked, in the speech of the parents (and in the speech of *their* parents, etc.); on our account, by contrast, the frequency distribution of locatives is something which may be explained rather than an explanation in itself.

In the six experiments that follow, we test the hypothesis that the universal of Object Affectedness is used by children to predict the syntactic privileges of verbs, but that children must learn what counts as affected. From our proposal it follows that children will be productive--and in fact, before they figure out what counts as affected they will overgenerate locative forms. Moreover, we predict that instances of overgeneration should be associated with corresponding misinterpretations of verb meaning. In the first two experiments, we make use of picture sets to test the child's understanding of fine-grained aspects of verb meaning--aspects which must be tested in a controlled, experimental setting. In designing these studies, we have paid special attention to the source of semantic errors that might apply to verbs such as *fill* and *pour*, or *empty* and *dump*. In particular, we have assumed that a child might misinterpret *fill* to specify a pouring manner, or *pour* to specify a full endstate; *empty* to specify a dumping manner, or *dump* to specify an empty endstate. Our rationale is simply that children are likely to be exposed to many events which are both pouring and filling (since pouring is a common means of filling), or both dumping and emptying (since dumping is a common means of emptying). We also assess each child's willingness to produce both content- and container-locative forms. By thus testing for the syntax and semantics of particular verbs *within child*, we can address three questions. First, will children overgenerate locative forms, producing sentences like *the man is filling the water*? Second, will children misinterpret the meanings of

locative verbs, perhaps thinking that (e.g.) *fill* specifies the particular manner in which a substance changes location instead of the particular change of state that a container undergoes? Third, and most crucially, will syntactic and semantic errors be associated--for example, will children who misinterpret the meaning of *fill* to specify the manner in which a substance changes location be more inclined to encode the argument corresponding to that substance as the direct object? If the answer to each of these questions is *yes*, we will have strong support for the hypothesis that verb meaning and syntax are linked in the lexicons of language learners.

In the final four experiments, we again test for the correspondence between verb syntax and semantics, but in a more direct fashion. In these experiments, we manipulate verb meaning as an independent variable by teaching subjects made-up verbs (in a neutral syntax; e.g., *this is mooping*) that differ in whether the content or the container is affected in a particular and salient way. We then test each subject's willingness to produce locative forms of the made-up verbs. For verbs in which the content is affected, we predict that subjects should produce relatively more content-locative forms; for verbs in which the container is affected, we predict that subjects should produce relatively more container-locative forms. If these predictions are borne out, we have evidence that the meanings of verbs *must* be used under some circumstances to predict their syntactic privileges.

Experiment 1

In our initial experiment, we tested children on their ability to produce and understand six common locative verbs: *pour*, *fill*, *dump*, *empty*, *stuff*, and *splash*.

Method

Subjects. Sixty-four subjects, all native speakers of English living in the Boston area, participated in the study. Forty-eight of the subjects were children, falling into three age groups: sixteen between the ages of 2;6 (years;months) and 3;5 (mean 3;1); sixteen between the ages of 3;6 and 4;5 (mean 3;11); and sixteen between the ages of 4;6 and 5;11 (mean 5;0). The children were drawn from middle-class day-care and after-school programs in Cambridge, Needham, Newton, and Watertown. The remaining sixteen subjects were undergraduates at MIT, ranging from 18 to 22 years of age. The adult subjects were paid for their participation.

Eight subjects, all children, were replaced in design: seven children were replaced because of their unwillingness or inability to perform the production task; one child was replaced for failing to cooperate with the experimenters.

Materials. The materials for the study consisted of forty-nine line drawings. Each drawing was composed of two *panels*, much like a comic strip. The use of panels allowed us to separate each action into a sequence of two parts, and thus to tease apart which component of an action was taken by a subject to be essential to the meaning of a verb. Specifically, the first panel of each picture depicted the manner in which a substance changed location during the course of the action, while the second panel of each picture depicted the endstate of a container as a result of the action. For example (Figure 1), the first panel might show a woman in the process of pouring water from a pitcher into a glass; the second panel might show an empty glass next to a

Figure 1

Pouring -&- Spilling

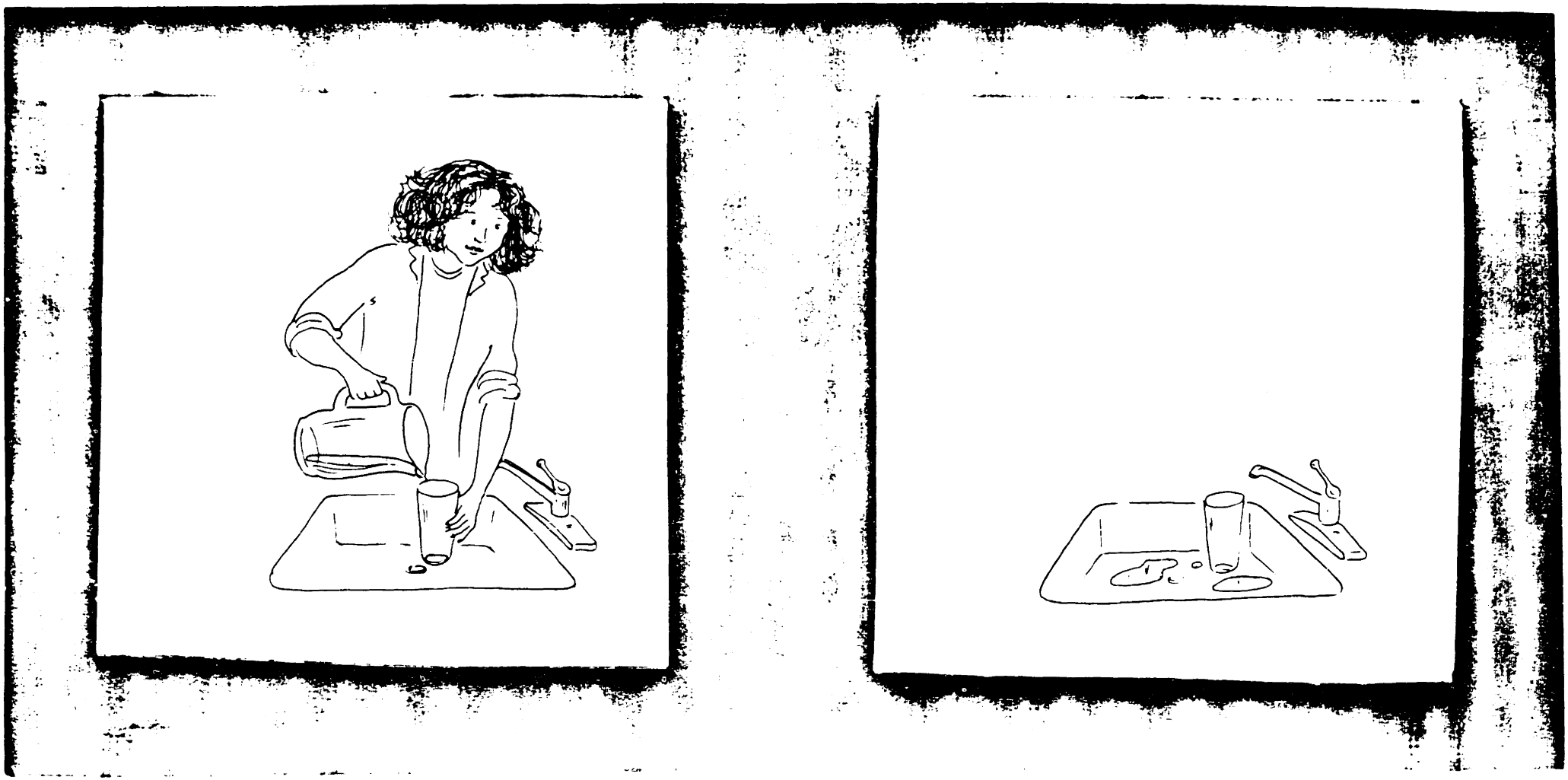
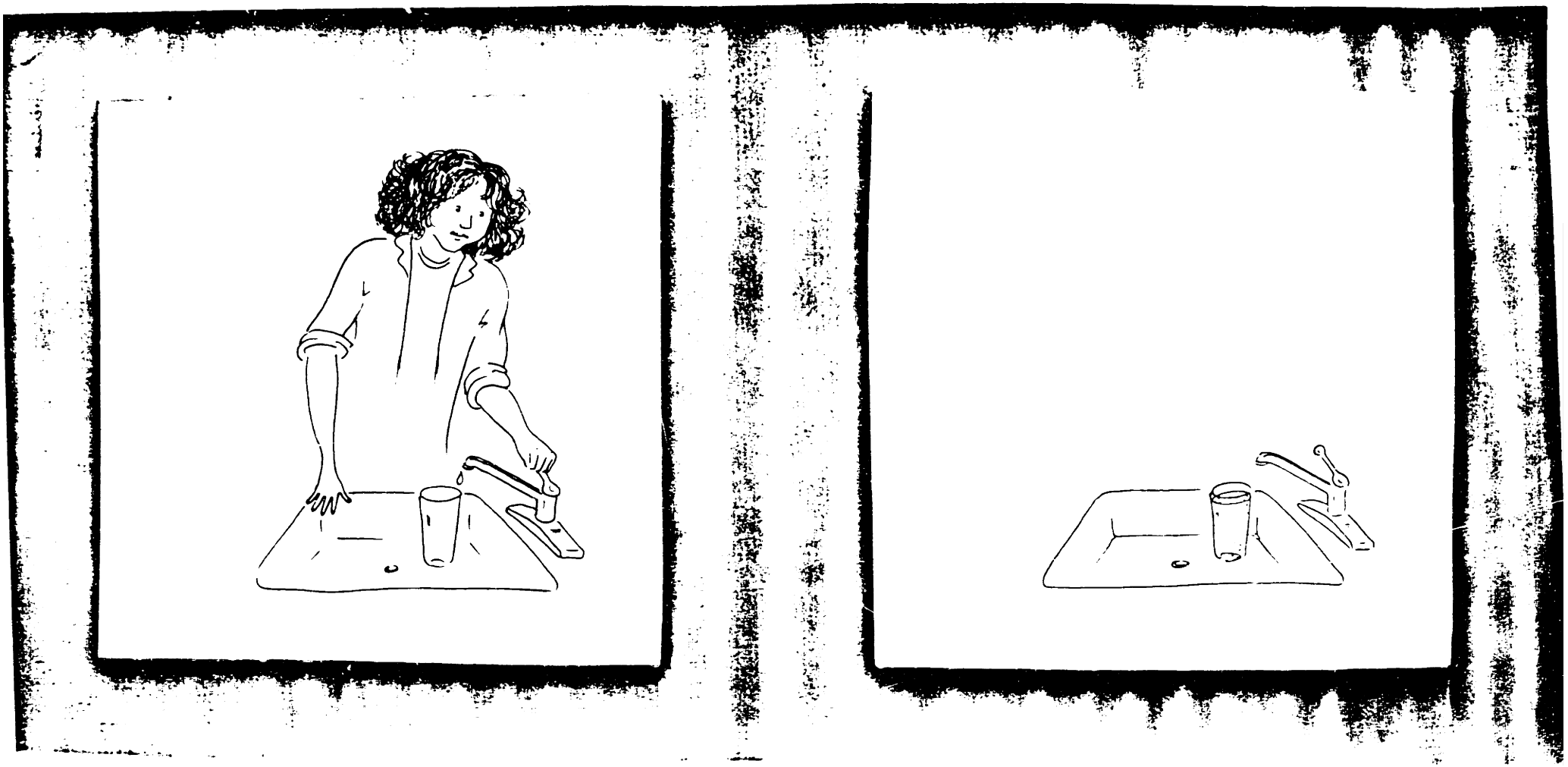


Figure 2

Dripping -&- Filling



puddle of water--the inference being that the woman has spilled the water. By contrast, in another picture (Figure 2), the first panel might show a woman in the process of turning on a faucet, allowing water to drip from a spigot into a glass; the second panel might show a glass full of water. A subject who knows that the meaning of *pour* has something essential to do with the manner in which a substance changes location should choose Figure 1 over Figure 2 as the better example of pouring. On the other hand, a subject who knows that the meaning of *fill* has something essential to do with the endstate of a container should choose Figure 2 over Figure 1 as the better example of filling.

Of the forty-nine drawings, one depicted a boy first hitting a ball with a bat (first panel), with the ball subsequently breaking a window (second panel). This picture was used to ensure that the subjects understood the format of the drawings--in particular, that the two states depicted in the panels were causally related to one another.

The remaining forty-eight drawings were used to test the ability of subjects to understand and produce six common locative verbs: *pour*, *fill*, *dump*, *empty*, *stuff*, and *splash*. In the comprehension task for a particular verb, subjects were forced to choose which of two pictures, differing in manner and/or endstate, best represented the meaning of the verb. The choice of drawings depicting pouring and filling, dumping and emptying was governed by the assumption that children might selectively confuse the interpretation of pairs of verbs; for example, a child might interpret *fill* to specify a pouring manner or *pour* to specify a full endstate; *empty* to specify a dumping manner or *dump* to specify an empty endstate. In addition, this choice of drawings and verbs provided us with a built-in control: because the verbs in these pairs are closely related in meaning, we were able to test subjects' interpretation of both verbs of a pair using the same sets of pictures (across subjects). This control helped to insure that subjects' responses were not due to the salience of the pictures themselves.

An example will clarify our use of picture sets for pairs of verbs. Every subject was shown the picture displayed in Figure 3. Some of the subjects were told "this is pouring"; others, "this is filling." We could do this because *pour* and *fill* have overlapping extensions and because the picture satisfies both extensions. After establishing a common frame of reference of what is pouring or what is filling, we then went on to test whether or not subjects knew the difference between the two verb meanings by presenting a forced choice between two nonoverlapping pictures--such as those in Figures 1 and 2. Notice that each of these pictures preserves one panel from the original (Figure 3) and introduces a new panel. Figure 1 preserves the panel depicting the manner in which a substance changes location, whereas Figure 2 preserves the panel depicting a change in the state of the container. We then asked, about these pictures, "Which one is pouring?" Or, for other subjects, "Which one is filling?" Because of our choice of new panels in these pictures, we can make the following claim about a subject who knows the difference between the verb meanings: when asked to choose which of these two pictures is pouring, he should choose the picture preserving the manner of the action (Figure 1); when asked to choose which of these two pictures is filling, he should choose the picture preserving the endstate of the action (Figure 2). Crucially, any systematic difference in response could only be due to a difference in verb meaning, and not to a difference in the pictures themselves.

Twelve pictures were shared for the verbs *pour* and *fill*. Of these twelve, subsets of three pictures concerned the same scenario--that is, the same human agent, the same potential container, and the same potential content. In each subset, one picture was always ambiguous between pouring and filling (e.g., Figure 3), one picture depicted pouring but not filling (Figure 1), and one picture depicted filling but not pouring (Figure 2). Similarly, a set of twelve pictures was shared for the verbs *dump* and *empty*. Table 2 lists descriptions of the drawings used in the testing of *pour/fill* and *dump/empty*, organized by scenario. Within child,

Figure 3

Pouring -&- Filling

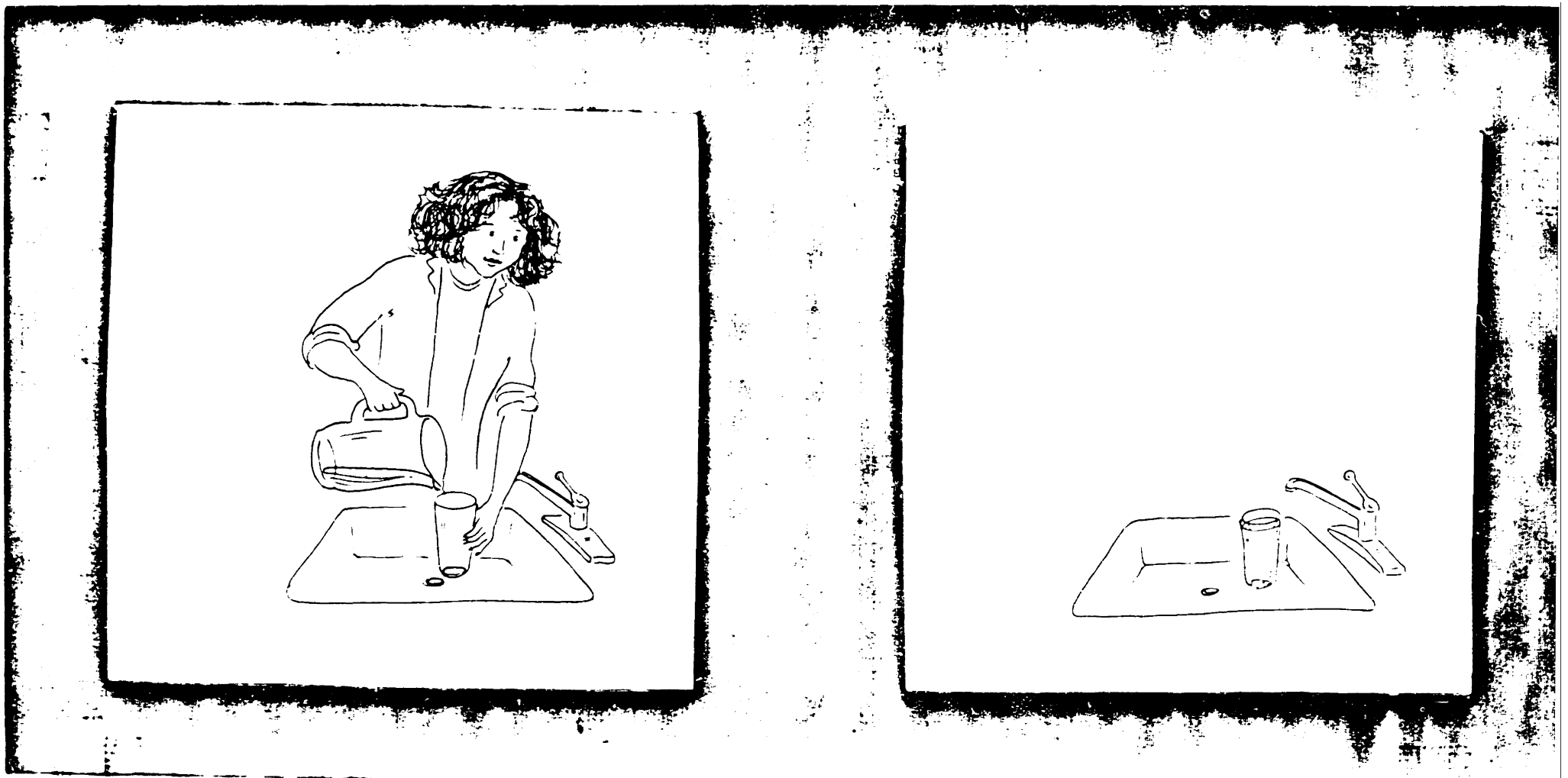


Table 2

Picture Sets Used in the Testing of *Pour/Fill* and *Dump/Empty*

PANEL 1 (MANNER)	PANEL 2 (ENDSTATE)
Scenario A1 (*pour/fill*)	
man pouring water from bucket into sink	sink filled with water
man pouring water from bucket into sink	empty sink/spilled water
man dripping water from faucet into sink	sink filled with water
Scenario A2 (*pour/fill*)	
girl pouring honey from bottle into bowl	bowl filled with honey
girl dripping honey from fork into bowl	bowl filled with honey
girl pouring honey from bottle into bowl	empty bowl/spilled honey
Scenario B1 (*pour/fill*)	
boy pouring paint from can into bucket	bucket filled with paint
boy dripping paint from brush into bucket	bucket filled with paint
boy pouring paint from can into bucket	empty bucket/spilled paint
Scenario B2 (*pour/fill*)	
woman pouring water from pitcher into glass	glass filled with water
woman pouring water from pitcher into glass	empty glass/spilled water
woman dripping water from faucet into glass	glass filled with water
Scenario C1 (*dump/empty*)	
man dumping ice cream from carton into bowl	empty carton/ice cream in bowl
man scooping ice cream from carton into bowl	empty carton/ice cream in bowl
man dumping ice cream from carton into bowl	1/2 empty carton/some ice cream in bowl
Scenario C2 (*dump/empty*)	
girl dumping playdo from can onto table	empty can/playdo on table
girl dumping playdo from can onto table	1/2 empty can/some playdo on table
girl scooping playdo from can onto table	empty can/playdo on table
Scenario D1 (*dump/empty*)	
woman dumping salad from bowl onto plate	empty bowl/salad on plate
woman dumping salad from bowl onto plate	1/2 empty bowl/some salad on plate
woman scooping salad from bowl onto plate	empty bowl/salad on plate
Scenario D2 (*dump/empty*)	
boy dumping sand from pail onto towel	empty pail/sand on towel
boy scooping sand from pail onto towel	empty pail/sand on towel
boy dumping sand from pail onto towel	1/2 empty pail/some sand on towel

Note: each line corresponds to a drawing composed of one manner panel and one endstate panel. For each subset of three drawings (e.g., A1), the first drawing was displayed *before* the remaining two; the second drawing was always displayed on the experimenter's right (the child's left); the third drawing was always displayed on the experimenter's left (the child's right).

two subsets (six pictures) were used for the testing of each verb; across children, each subset was used equally often in the testing of either verb of a pair. For the *pour/fill* sets, the manner distractor always depicted dripping and the endstate distractor always depicted an empty container (the contents were spilled); for the *dump/empty* sets, the manner distractor always depicted scooping and the endstate distractor always depicted a half-empty container.

The picture sets for the verbs *stuff* and *splash* differed from those above. Consider the verb *stuff*. Because the verb specifies both a characteristic motion and a characteristic endstate (according to our intuitions), we wanted to test a subject's sensitivity to each aspect of meaning separately. We did this by treating *stuff* as if it were two verbs--one having to do with each aspect of meaning. For example, a picture set for the "endstate" reading of *stuff* was composed of drawings which differed only in their endstate panels: first we showed each subject a picture such as that in Figure 4, saying "this is stuffing." Then, we presented two more pictures--Figures 5 and 6--asking "which of these is stuffing?" Notice that each of these pictures preserves the same panel of the original--in this case the panel depicting the manner in which a substance changes location. What differs between the two is the second panel. Using the original second panel (in Figure 4) as a foil, it is clear that Figure 5 (showing a *stuffed* container) is a better depiction of stuffing than Figure 6 (showing a half-full container). And the basis for this judgment has everything to do with our sensitivity to the change of state of the container, and nothing to do with the change in the location of its content.

Table 3 lists glosses of the twelve pictures used in the testing of *stuff* and the twelve pictures used in the testing of *splash*. Unlike the picture sets used for *pouring* and *filling*, or for *dumping* and *emptying*, none of these pictures is ambiguous in the depiction of the relevant locative events. Each set of twelve is divisible into two subsets of six pictures, with each

Figure 4

Dropping in the Contents -&- A Full Container (Foil)

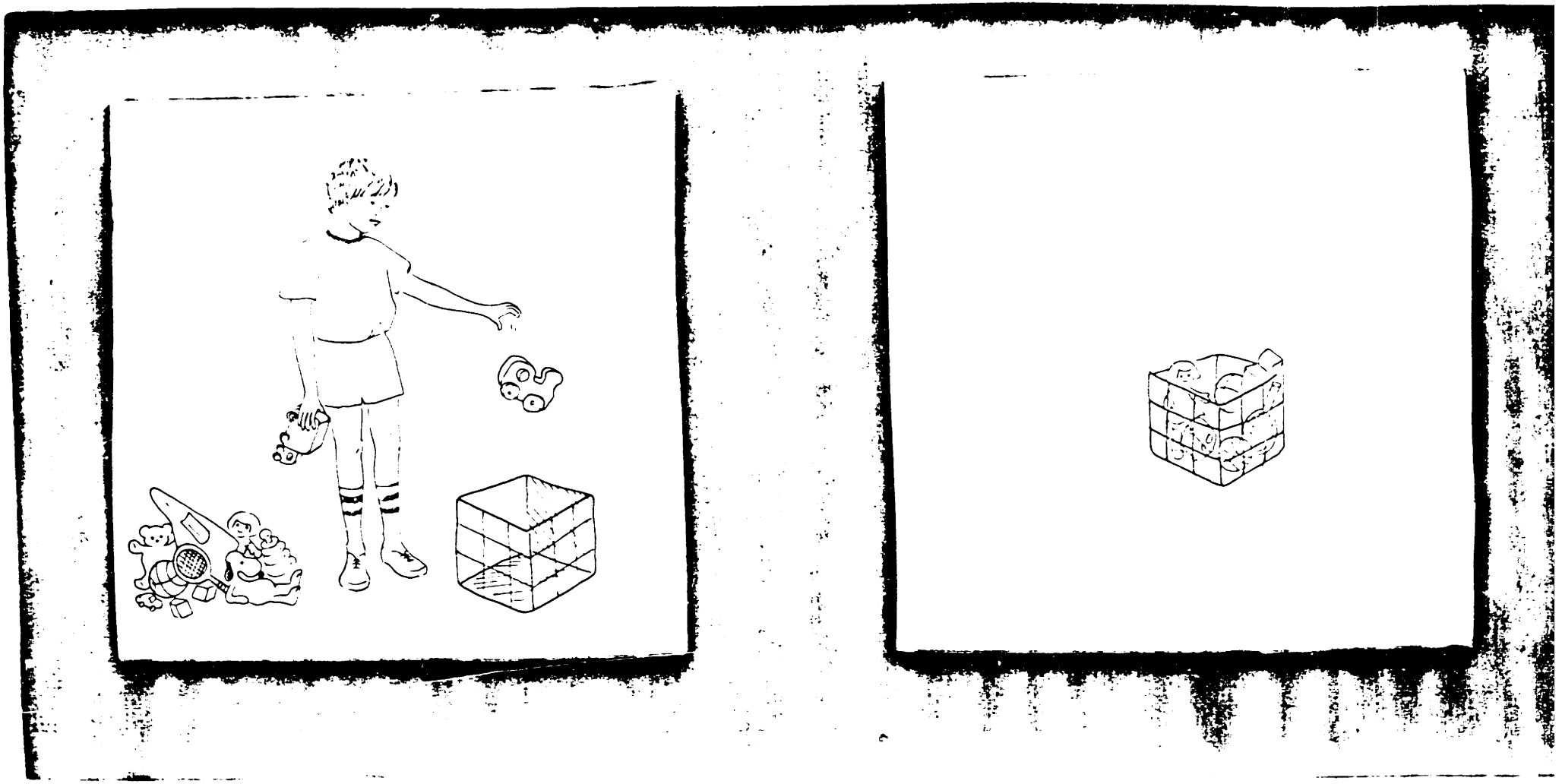


Figure 5

Dropping in the Contents -&- A Bulging Container

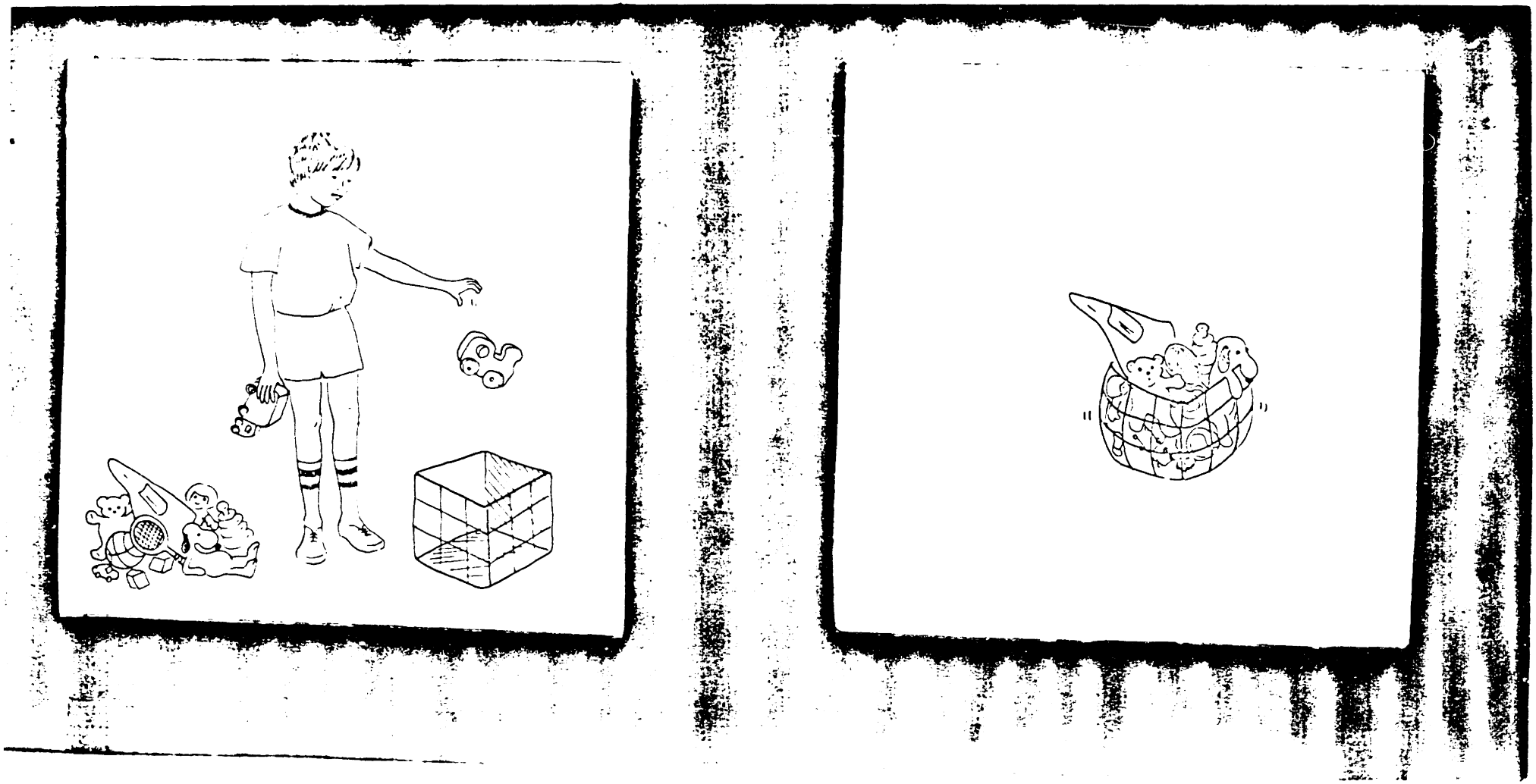


Figure 6

Dropping in the Contents -&- A Half-Full Container

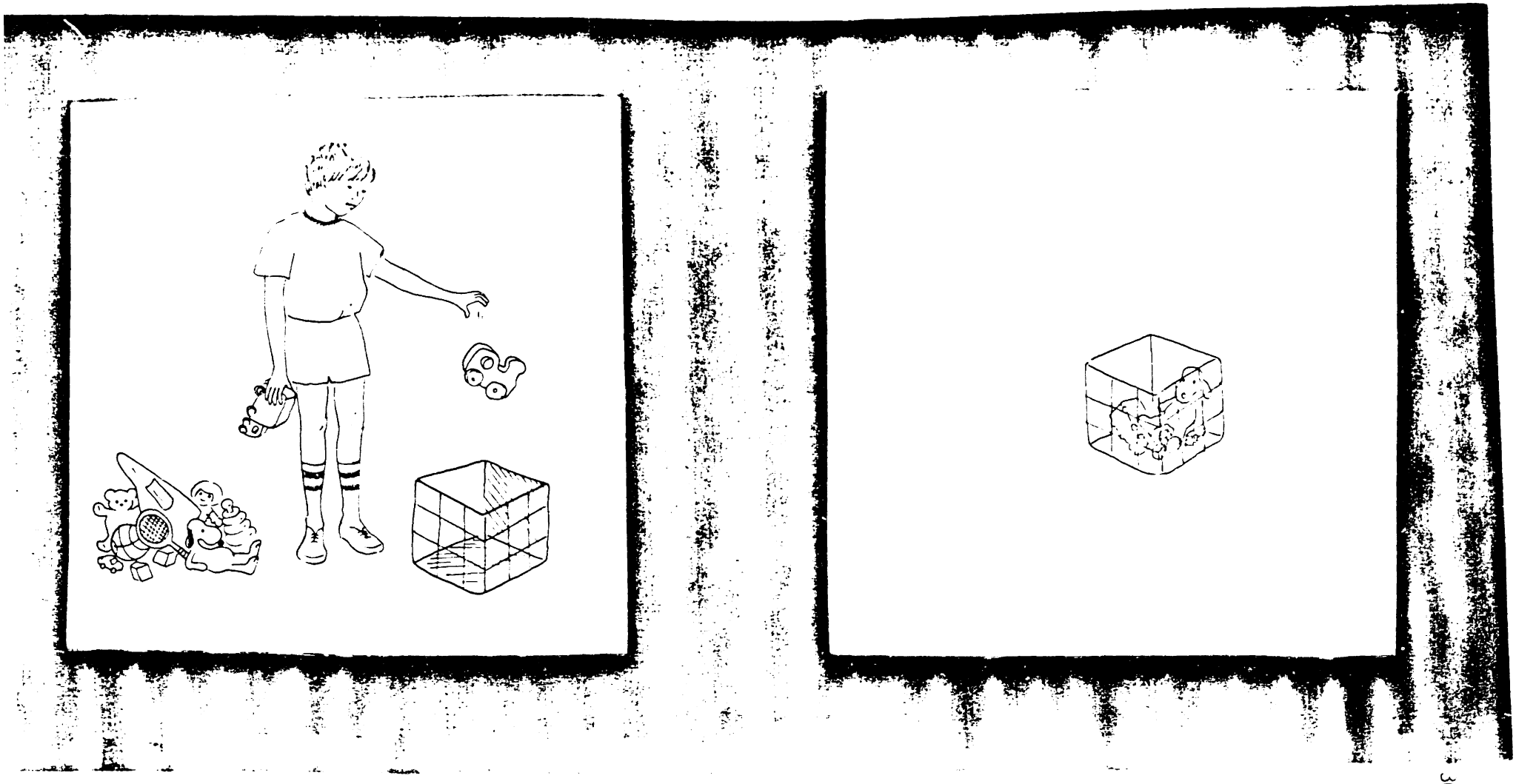


Table 3

Picture Sets Used in the Testing of *Stuff* and *Splash*

PANEL 1 (MANNER)	PANEL 2 (ENDSTATE)
Scenario E1 (*stuff*)	
woman dropping clothes into hamper	hamper filled with clothes
woman dropping clothes into hamper	hamper 1/2 filled with clothes
woman dropping clothes into hamper	hamper bulging with clothes
woman dropping clothes into hamper	hamper filled with clothes
woman forcing clothes into hamper	hamper filled with clothes
woman scooping up clothes with hamper	hamper filled with clothes
Scenario E2 (*stuff*)	
boy dropping toys into box	box filled with toys
boy scooping up toys with box	box filled with toys
boy forcing toys into box	box filled with toys
boy dropping toys into box	box filled with toys
boy dropping toys into box	box bulging with toys
boy dropping toys into box	box 1/2 filled with toys
Scenario F1 (*splash*)	
boy splashing water with feet from pool onto girl	damp girl
boy splashing water with hands fr. pool onto girl	damp girl
boy pushing girl into pool	damp girl
boy splashing water with feet from pool onto girl	damp girl
boy splashing water with feet from pool onto girl	dry girl
boy splashing water with feet from pool onto girl	drenched girl
Scenario F2 (*splash*)	
girl splashing water with feet from pool onto boy	damp boy
girl splashing water with feet from pool onto boy	drenched boy
girl splashing water with feet from pool onto boy	dry boy
girl splashing water with feet from pool onto boy	damp boy
girl pushing boy into pool	damp boy
girl splashing water with hands from pool onto boy	damp boy

Note: each line corresponds to a drawing composed of one manner panel and one endstate panel. For each subset of three drawings, the first drawing was displayed *before* the remaining two; the second drawing was always displayed on the experimenter's right (the child's left); the third drawing was always displayed on the experimenter's left (the child's right).

subset of six concerning a particular scenario; each subset of six is further divisible into two subsets of three pictures, with each subset of three differing in one panel (e.g., manner) but not the other (e.g., endstate). The following panels were used in testing for sensitivity to the manner of splashing (splashing with feet (foil), splashing with hands, pushing with hands); the endstate of splashing (a damp goal (foil), a wet goal, a dry goal); the manner of stuffing (dropping in the contents (foil), forcing in the contents, scooping up the contents with the container); the endstate of stuffing (a full container (foil), a bulging container, a half-full container).

Part of the reason for using two different types of picture sets was exploratory; it was a new method, and we wanted to test the waters. But let's be explicit about the differences between them. The picture sets used for *pour/fill* and *dump/empty* were designed to assess a *bias* in the interpretation of the verb's meaning. Specifically, alternative pictures in the forced-choice task differed from each other in *both* panels (and from the original picture in one panel), so that subjects were always choosing between a depiction of (e.g.) pouring, but not filling (e.g., Figure 1) and a depiction of filling, but not pouring (e.g., Figure 2). Notice that what these picture sets *don't* provide is a pure measure of *sensitivity* to a particular type of meaning; bias towards a component of meaning necessitates sensitivity to that component, but not the converse. As a consequence, if a child prefers the picture which depicts the manner in which a substance moves, it doesn't mean that the child is insensitive to the endstate of the container. For example, *filling* may be judged as having more to do with the pouring motion than the full endstate, but it could be the case that both meanings are, for that subject, essential to filling. But because the subject is forced to choose, we lose some information. The picture sets for *splash* and *stuff*, on the other hand, were designed to assess *sensitivity* to a particular component of meaning. For these sets, alternative pictures in the forced-choice task differed

from each other in *one* panel (and from the original picture in the same panel), so that subjects were always choosing between two depictions of manner or two depictions of endstate (e.g., Figures 5 and 6). An obvious limitation of the sensitivity test, but not necessarily the bias test, is its inability to distinguish properties of verb meaning that are essential from those that are merely typical. Thus, subjects never had to decide which of two components of meaning--manner or endstate--was more important, and perhaps essential, to the meaning of *splash* or *stuff*.

A further note about the drawings is that they provided for not only a *constant* depiction of verb meanings across trials, but also a *simultaneous* presentation of alternatives in the forced choice between possible verb meanings. In this way, we eliminated any confound between choice and temporal order of presentation which would have obtained had we acted out the alternative verb meanings. In addition, the drawings themselves were rendered in black ink, with a depiction of the container and its (potential) content occurring in each panel. Container and content were thus given equal representation in the drawings.

Finally, we wanted to avoid a response bias whereby subjects would prefer one panel to the other simply because verbs specifying manner outnumbered verbs specifying endstate, or vice versa. According to our own semantic analysis, the verbs *pour* and *dump* specify the way in which a substance changes location, whereas *fill* and *empty* are concerned with the properties of a container--the goal of motion in the case of *fill*, and the source of motion in the case of *empty*. The verbs *splash* and *stuff*, we thought, could specify a manner and/or an endstate. A second response bias which we tried to preempt has to do with the syntactic frames in which adults accept these verbs: *pour* and *dump* take only the content as direct object; *fill* and *empty* take the container as direct object; and *splash* and *stuff* are alternators. (To anticipate our findings, we note that *empty* is actually an alternator: most English speakers accept the sentence

John emptied the carton of ice cream; and some of them will also accept the sentence *John emptied ice cream from the carton*. Our intuition is that the latter sentence, with the content as direct object, isn't as acceptable as the first (cf., *John emptied the carton*; **John emptied the ice cream*), and our guess is that no significant response bias was introduced by testing *empty*.)

Procedure. The procedure consisted of assessing each subject's ability to comprehend and produce six locative verbs. Each adult was tested individually in a single session. Due to the length of the procedure, however, each child was tested individually in two half-hour sessions. The two sessions were separated by as few as 2 days to as many as 42 days, with a mean separation of 11.3 days. Subjects were tested in an area which was as free as possible of potential distractions (e.g., other children). For the children, the comprehension and production tasks were introduced as games, the object of which was to teach English words to puppets. Furthermore, two experimenters participated in testing each child: one experimenter engaged the child and elicited the responses; the second experimenter observed the task and recorded the responses.

Before testing their knowledge of locative verbs, we introduced subjects to the format of the pictures by presenting them with the drawing of a boy hitting a ball with a bat (first panel), with the ball subsequently breaking a window (second panel). Specifically, subjects were first asked to describe each panel separately, and then both panels together. Our goal was to insure that subjects ultimately interpreted both panels as part of a coherent drawing. Towards this goal, the experimenter modeled the sentence *the boy is breaking the window* if subjects did not spontaneously use an appropriate causative verb (e.g., *break*, *smash*) to describe the complete drawing.

During the main body of the experimental session, we tested each verb one at a time. As described above, we used sets of three pictures, first to assess verb meaning and then to

prompt for productions. After a child had chosen one picture or the other as depicting pouring, for example, the chosen picture was then used to elicit productions of *pour*. Furthermore, each cycle of comprehension and production trials was performed twice per verb or verb reading, the second time with a new set of three pictures. (Notice therefore that twice as many comprehension and production trials were conducted for each of *stuff* and *spill* as for each of the other verbs.) The order of verbs itself was balanced across subjects so that verbs with overlapping extensions (e.g., *pour* and *fill*) were never tested in consecutive order. In fact, to lessen the possibility that children would (artifactually) confuse verbs with overlapping extensions, such verbs were never tested in the same session.

In testing the ability of subjects to comprehend the meaning of a verb (e.g., *fill*), the experimenter began by introducing the "constituents" of the first drawing in a set, and by explicitly labelling the drawing as a depiction of filling. For example (Figure 3), the experimenter would say to a subject "Look at the first picture (panel): there's a woman, a pitcher, water, and a glass. Look at the second picture (panel): there's the glass and the water. Now look at both pictures: when the woman does *this* (experimenter pointing to first panel), it ends up like *that* (experimenter pointing to second panel). And it's called *filling*. *This* (experimenter gesturing with entire drawing) is *filling*." The experimenter would then remove the first drawing and present the two alternative drawings in the forced choice (e.g., Figures 1 and 2). The constituents in each drawing would be introduced, as above, starting with the drawing on the experimenter's right. Neither of these drawings, of course, would be labelled as a depiction of filling. Instead, the experimenter would ask, "Which of *these* (experimenter gesturing with both drawings) is *filling*?" If a subject did not clearly indicate either one drawing or the other, the experimenter repeated the question.

As described earlier, two subsets (six pictures) were used for the testing of each verb or verb reading. Across subjects, each subset was used equally often in the testing of either verb of a pair. Furthermore, the position of alternative pictures in the forced choices was balanced within subject so that, for the two forced choices per verb or verb reading, a given type of picture (e.g., 'pouring -&- emptying') appeared on the right as often as it did on the left.

In testing the ability of subjects to produce a verb (e.g., *fill*) in both locative forms, we did more than simply ask the child to describe the picture that he or she had chosen in the preceding comprehension task. In particular, we posed one of two queries--either a *content-topic* query, which focuses the content (e.g., water), or a *container-topic* query, which focuses the container (e.g., the glass). Thus, if the chosen picture corresponds to Figure 2, the experimenter would focus the container in the following way: "Point to the glass; say *glass* ...(experimenter waits for response); say *filling* ...(experimenter waits for response); what is the woman doing to the *glass*?" Notice that a natural response to this question is "She's filling it with water"--where the container is encoded as direct object. In the same way, a *content-topic* query will set up a discourse context favoring a locative response with the content as direct object. Of course, whether or not a given form is uttered depends upon what the subject takes to be a *possible* syntactic expression for a particular verb. Thus, we used the pragmatics of the query in order to flush out the *range* of possible locative forms that a verb can take. Accordingly, each verb or verb reading was tested with both types of query (once after each comprehension trial). (For a fuller discussion of the same methodology applied to eliciting datives, see Gropen, Pinker, Hollander, Goldberg, and Wilson, 1989.) Furthermore, in those trials where a subject failed to indicate an unambiguous direct object, we followed up with a secondary prompt: "filling what?", or "filling ___?" where a completion by the subject is the contextually appropriate response. (Note that this alternative query was always used instead of "splashing what?" in the trials for *splash*, where the content and container differed in animacy;

otherwise, only "(splashing) water" would have been the appropriate response.) And if the secondary query didn't work, we asked, "Is the woman filling the glass or filling the water?" (with order of disjuncts balanced within verb). In this way, we scored three levels of response in the production task.

Because the topic of the queries in the production trials has a direct influence on the types of productions elicited, we balanced the order of query topic used in the testing of each verb across the subjects in an age group. In fact, verb testing order, picture set combinations, and query order were counterbalanced across subjects within each age group.

Scoring. In the comprehension task, responses were scored according to whether the chosen picture was consistent or inconsistent in manner or endstate with the meaning of the verb. In the case of the bias tests (for *pour*, *fill*, *dump*, and *empty*), the criterion of consistency was equivalent to deciding whether the chosen picture preserved the manner or endstate of the original drawing. In the case of the sensitivity tests (for *stuff* and *splash*), this criterion was equivalent to deciding whether the contrasted panel of the chosen picture provided a good match or a poor match to adult intuitions. Instead of presuming that we knew the meanings of the verbs, however, we viewed the adult comprehension (and production) data as the standard: the final arbiters of verb knowledge were the adult subjects in the experiment.

In the production task, responses were scored according to whether the direct object of an acceptable locative form corresponded to the content or the container in the described picture. An acceptable locative form was one in which the child or adult indicated both the appropriate verb and an unambiguous direct object. The level of locative response was also scored: whether the subject responded to the first query, the second query (e.g., "filling what?"), or the third query (e.g., "filling the glass or filling the water?"). Responses which were undecipherable or not clearly locative (e.g., intransitive responses such as *she's pouring with*

the water) were coded as *other* . The experimenter also coded additional information in the responses, including the utterance of oblique objects (e.g., *the cup* in *he poured water into the cup*) and particles (e.g., *out* in *he poured water out*).

Because we used queries which explicitly focused either the container or the content, it was typical for subjects to respond with utterances containing a pronominal reference. For example, to the query, "What is the woman doing to the glass?" a subject might respond, "She's filling it." Although this is an acceptable response *in discourse*, it does not satisfy our criteria for an unambiguous experimental response. In particular, we did not assume that subjects--especially children--were fastidious in their use of *it* to refer to the previously focused entity. Instead, pronominal reference to a content or container was counted as acceptable (at the first level of response) if: a) the referent of the pronoun was disambiguated by the presence of an oblique object or particle (e.g., *it* in *she poured it into the glass* or *she poured it out* can only refer to the water); or b) the referent of the pronoun was disambiguated by the plurality or gender of the pronoun (i.e., *them*, *he*, and *she* were considered unambiguous, versus the unmarked *it*; the pronoun *it* was considered unambiguous only in the trials with *splash*, where the content (water) and container (a boy or girl) differed in animacy); or c) the reference could be subsequently tied down via the second query.

Design. For the comprehension task, we employed a 6 x 3 factorial design with the within-subject factor of Verb (*pour* vs. *fill* vs. *dump* vs. *empty* vs. *stuff* vs *splash*) and the between-subjects factor of Age Group (2;6-3;5 vs. 3;6-4;5 vs. 4;6-5;11 vs. adult). The dependent variable was the proportion of trials in which the chosen picture was consistent or inconsistent in manner or endstate with the meaning of the verb. The adults' performance in the comprehension task was regarded as the standard for the purposes of establishing the meanings of the verbs.

For the production task, we employed a 2 x 6 x 3 factorial design with the within-subject factors of Query Topic (content vs. container) and Verb (*pour* vs. *fill* vs. *dump* vs. *empty* vs. *stuff* vs *splash*), and the between-subjects factor of Age Group (2;6-3;5 vs 3;6-4;5 vs. 4;6-5;11 vs. adult). The dependent variable was the proportion of trials in which either the content or container was encoded as direct object. As above, the adults' performance in the production task was regarded normatively for the purposes of deriving the proportion of trials in which children produced non-standard ("incorrect") forms--sentences which could not have been heard in the positive input, but were nonetheless uttered by children.

Results and Discussion

Our principal findings were that children overgenerated the locatives of *fill* and *empty* with the content encoded as direct object (e.g., *fill the water*; *empty the playdo*); that children misinterpreted the meaning of *fill* and *empty* as having something essential to do with the manner with which a substance changes location; and that instances of overgeneration with *fill* and (especially) *empty* were associated with corresponding misinterpretations of verb meaning. In elaborating on these results below, we'll look at production first, then comprehension, and finally the association of verb syntax and verb meaning.

Production. On the issue of production, we will focus on the occurrence or non-occurrence of *incorrect* locative forms. We will therefore not be primarily concerned with alternating verbs, which occur in both forms by definition. With regard to non-alternating verbs, then, is it the case that children will utter locative forms of *fill* or *empty* with the content, and not the container, encoded as direct object? Or locative forms of *pour* or *dump* with the container, and not the content, encoded as direct object? Before we can answer these questions, however, we must first confirm or disconfirm our intuitions about the non-alternating status of these verbs. In Table 4 we've presented the proportion of trials in which subjects produced content- and

Table 4

Proportion of Trials in which Content- and Container-Locatives of Each Verb were Produced as a Function of Query Topic and Age Group

	AGE GROUP			
	2;6-3;5	3;6-4;5	4;6-5;11	Adult
VERB-FORM				
Pour				
Content Locatives				
Content-Topic Query	0.94	1.00	1.00	1.00
Container-Topic Query	0.88	1.00	1.00	1.00
Mean	0.91 (18/10/1)	1.00 (29/3/0)	1.00 (29/3/0)	1.00
Container Locatives				
Content-Topic Query	0.06	0.00	0.00	0.00
Container-Topic Query	0.12	0.00	0.00	0.00
Mean	0.09 (1/2/0)	0.00	0.00	0.00
Fill				
Content Locatives				
Content-Topic Query	0.56	0.62	0.19	0.06
Container-Topic Query	0.50	0.44	0.50	0.00
Mean	0.53 (9/8/0)	0.53 (9/7/1)	0.34 (5/6/0)	0.03
Container Locatives				
Content-Topic Query	0.44	0.31	0.75	0.94
Container-Topic Query	0.50	0.56	0.44	0.94
Mean	0.47 (4/10/1)	0.44 (10/1/3)	0.59 (9/9/1)	0.94
Dump				
Content Locatives				
Content-Topic Query	1.00	1.00	1.00	1.00
Container-Topic Query	0.88	0.81	0.94	0.88
Mean	0.94 (19/10/1)	0.91 (27/2/0)	0.97 (26/4/1)	0.94
Container Locatives				
Content-Topic Query	0.00	0.00	0.00	0.00
Container-Topic Query	0.12	0.19	0.06	0.12
Mean	0.06 (2/0/0)	0.09 (1/1/1)	0.03 (1/0/0)	0.06

Note: The numerals in parentheses correspond to the frequencies of locatives produced at the 1°/2°/3° level of response. Adults always responded to the primary (1°) query.

Table 4 (continued)

Proportion of Trials in which Content- and Container-Locatives of Each Verb were Produced as a Function of Query Topic and Age Group

	AGE GROUP			
	2;6-3;5	3;6-4;5	4;6-5;11	Adult
VERB-FORM				
Empty				
Content Locatives				
Content-Topic Query	0.69	0.69	0.69	0.75
Container-Topic Query	0.50	0.56	0.50	0.25
Mean	0.59 (6/8/5)	0.62 (12/7/1)	0.59 (13/6/0)	0.50
Container Locatives				
Content-Topic Query	0.31	0.31	0.31	0.25
Container-Topic Query	0.44	0.38	0.50	0.75
Mean	0.38 (5/5/2)	0.34 (4/4/3)	0.41 (11/1/1)	0.50
Stuff				
Content Locatives				
Content-Topic Query	0.88	0.91	0.91	0.91
Container-Topic Query	0.81	0.69	0.59	0.56
Mean	0.84 (17/34/3)	0.80 (30/19/2)	0.75 (36/11/1)	0.73
Container Locatives				
Content-Topic Query	0.09	0.09	0.09	0.09
Container-Topic Query	0.19	0.31	0.41	0.44
Mean	0.14 (2/7/0)	0.20 (8/3/2)	0.25 (12/3/1)	0.27
Splash				
Content Locatives				
Content-Topic Query	0.59	0.47	0.72	0.81
Container-Topic Query	0.56	0.44	0.25	0.25
Mean	0.58 (22/9/6)	0.45 (23/5/1)	0.48 (29/1/1)	0.53
Container Locatives				
Content-Topic Query	0.31	0.47	0.22	0.16
Container-Topic Query	0.34	0.56	0.72	0.72
Mean	0.33 (15/1/5)	0.52 (23/2/8)	0.47 (25/2/3)	0.44

Note: The numerals in parentheses correspond to the frequencies of locatives produced at the 1°/2°/3° level of response. Adults always responded to the primary (1°) query.

container-locatives of each verb, as a function of query topic and age group. As should be clear, the performance by adults confirms our syntactic judgments about *fill*, *pour*, *dump*, *stuff*, and *splash*; in particular, *pour*, *fill*, and *dump* are non-alternators (we find a lack of *fill*-content, *pour*-container, and *dump*-container utterances); *stuff* and *splash* are alternators. *Empty*, contrary to our expectations, also turns out to be an alternator, with no apparent preference for either form.

Viewing the adult pattern of results as the standard, we can look for deviations from the standard in the production of non-alternators by the children. Specifically, the non-standard forms which could be produced by children, in principle, are of three types: *fill*-content, *pour*-container, or *dump*-container. Of these three forms, however, children were much more likely to produce the non-standard *fill* form than the other two: 30 children out of 48 produced at least one *fill*-content form (11, 11, and 8 children from young, mid, and old child groups, respectively), whereas only 2 children produced at least one *pour*-container form, and only 6 children produced at least one *dump*-container form. For each form, we performed an Analyses of Variance on the mean proportion of production trials in which the form was produced (i.e., the mean proportion of queries to which a particular form was given in response). The within-subject factor was Query Topic, and the between-subjects factor was Age Group. As expected, we found a significant main effect for age group in the ANOVA for the *fill*-content form, indicating that there is a significant difference, across age groups ($M_{young} = 0.53$, $M_{mid} = 0.53$, $M_{old} = 0.34$, $M_{adt} = 0.03$), in the mean proportion of queries to which the *fill*-content form was produced ($F(3, 60) = 6.63$, $p < .001$). No main effects for age group were found for the production of *dump*-container or *pour*-container forms.

We also found several effects which demonstrate the efficacy of manipulating query topic in order to encourage the production of both container and content locatives. First, we found a

main effect of query topic for the production of *dump*-container forms, indicating that more container locatives were produced in response to the container-topic query ($M = 0.12$; 8 forms out of 128 trials) than in response to the content-topic query ($M = 0.00$), $F(1, 60) = 8.73$, $p < .005$. This main effect clearly demonstrates the predictable influence of discourse topic on the subjects' choice of locative form; to the extent that *any* non-standard container locatives were uttered, they were produced in a context in which the role of the container was the discourse topic. Second, we found an interaction between age group and query topic for the production of *fill*-content forms ($F(3, 60) = 3.27$, $p < .05$): for the oldest children, but not for any other group, subjects produced more content locatives in response to the container-topic query than in response to the content-topic query. A post-hoc test showed that the oldest kids produced significantly more *fill*-content forms in reply to the container-topic query ($M = 0.50$) than in reply to the content-topic query ($M = 0.18$), $t(15) = 2.61$, two-tailed $p < .02$. We have no explanation for why the oldest children flouted the discourse function of locative forms; however, we note that, for every group of subjects except the oldest child group, the query which treated the identity of the content as old information *was* more successful in eliciting content-locative forms (though not significantly so according to post-hoc *t*-tests).

To address the issue of just which child groups deviated from the standard of adult locative production for non-alternating verbs, a series of planned one-tailed *t*-tests was performed on the difference, between child and adult groups, in the mean proportion of queries to which the non-standard forms were produced. What we find is that between the adult ($M = 0.03$) and the combined child groups ($M = 0.47$) there is a significant difference in the proportion of trials in which the *fill*-content form was produced ($t(62) = 4.11$, $p < .001$), but not in the proportion of trials in which the *pour*-container or *dump*-container forms were produced. Furthermore, the significant difference in mean proportion of *fill*-content production was upheld for each of the child groups analyzed separately against the adult group (from youngest to oldest, $t(30) =$

4.50, $p < .001$; $t(30) = 4.50$, $p < .001$; $t(30) = 3.01$, $p < .003$). We note here that the oldest children produced fewer *fill*-content forms (11 utterances, $M_{old} = 0.34$) than the younger children (17 utterances for each of the younger groups, $M = 0.53$), presumably because the oldest children have had more exposure to *fill*-container forms than the younger children. A post-hoc comparison revealed that this difference was not significant (two-tailed $t(46) = 1.48$, $p \approx .14$).

If children as young as 2;6-3;5 are producing *fill*-content forms, as our results show, what are we to make of Bowerman's (1982) observation that children younger than four years of age are accurate in choosing standard locative forms in their spontaneous speech? We suggest that part of the discrepancy may be due the efficacy of our production task in uncovering true linguistic capacity; in particular, by controlling the salience of contents and containers, and by manipulating the topic of our queries, we have encouraged children to utter rule-governed locative forms that they otherwise might not have been willing to produce. Indeed, the fact that young children are unwilling to produce *fill*-content forms spontaneously may reflect a conservatism in linguistic *behavior*, not linguistic capacity--fostered by the positive input of *fill*-container forms that children receive (as well as the free will that they possess). For these reasons, we suggest that the 'U'-shaped curve that Bowerman documents may reflect, in the case of locative verbs, the advent of linguistic risk-taking as well as the advent of reorganization.

On the other hand, given the discrepancies between induced and spontaneous production and the marked difference between adult and child performance, can we safely assume that the production data reflect the syntactic knowledge that subjects have of particular verbs, especially of the verb *fill*? One potential basis for concern is that, by design, we gave each subject the opportunity to respond at one or more of three levels of response. The issue here, in particular,

is whether or not the responses to the second and third queries are true reflections of syntactic knowledge; alternatively, these responses could be based on the relative salience of the content or container. Setting aside the fact that many (17/41, or 41%) of the responses to the second query (in *fill* trials) included some indication that they were syntactic in nature (i.e., the verb, a particle, and/or a prepositional phrase was uttered along with the direct object), we can separately analyze the *fill*-content responses to the *primary* query, which were clearly syntactic in nature. (See Table 4 for the frequency of locative production by level of response.) What we find is the same result reported above: there was a significant difference, between the adult ($M = 0.03$) and the combined child groups ($M = 0.24$), in the mean proportion of primary queries to which the *fill*-content form was produced (one-tailed $t(62) = 2.49$, $p < .01$). Furthermore, this result held true for each of the child groups ($M_{young} = 0.28$; $M_{mid} = 0.28$; $M_{old} = 0.16$) analyzed separately against the adult group (from youngest to oldest, one-tailed $t(30) = 2.60$, $p < .01$; one-tailed $t(30) = 2.60$, $p < .01$; one-tailed $t(30) = 1.85$, $p < .05$).

One other deviation from standard adult production is apparent in the child data for *fill*. Ordinarily, adults use the particle *up* in *fill*-container forms (e.g., *she filled the glass up*) to emphasize the completeness with which a container changes state (see Talmy, 1985; Fraser, 1971; Moravcsik, 1978). Some children, however, appear to use the particle *up* in a literal fashion, perhaps to indicate that the level of content in a container has risen as the result of the filling action (cf., *she filled the water up*). Specifically, children in the production task for *fill* used particles in 22 trials (out of 96); of these 22 particles, 18 were instances of *up*; of these 18 instances, 9 occurred in *fill*-container forms and 9 occurred in *fill*-content forms. The fact that children were equally likely to use *up* in either locative form turns on the ambiguity of the particle, and bears on the issue of how children interpret the *fill*-content form. We shall return to this issue in our discussion of the association between verb meaning and verb syntax.

With regard to the alternating verbs--*stuff*, *splash*, and *empty*--we can also look at children's deviations from standard adult production. In these cases, of course, the deviations may simply be a reflection of the input that children receive, rather than the result of a hypothesized universal of language. We will make the assumption, however, that the adult pattern of production is a rough correlate of the input that children typically receive; under this assumption, we can at least ascertain whether children have any *strong* non-standard preferences for one locative form or the other.

Before we look more carefully at differences between age groups, let us emphasize that we are *not* interested here in the occurrence or non-occurrence of incorrect forms. Instead, we shall use a measure of the *preference* of one form over the other. Specifically, the *preference score*--for a given subject--will correspond to the proportion of trials in which a content locative is produced *minus* the proportion of trials in which a container locative is produced. Preference scores will therefore range from +1.0 (a strong preference for content locatives) to -1.0 (a strong preference for container locatives). Note that our production data reflect preference for one locative form or the other only insofar as the production task is a crude forced-choice procedure (on the basis of the wording of the queries, in general, and the query topic, in particular). Thus, the definition of a preference score assumes that, for example, if children produce *fewer* container locative forms in the task than do adults, children must also be producing *more* content locatives than do adults. This assumption, and accordingly the notion of a preference score, are justified by the low mean proportion of queries (in alternator trials) to which *other* responses were produced (M = 16/640 or 0.025). (The mean proportion of *other* responses drops to three responses out of 384 (0.008) if one considers just the *empty* and *stuff* trials; the relatively high proportion of non-locative *splash* forms (13/256 or 0.051) is due in part to the acceptability of intransitive forms such as *the girl is splashing* (*with water/at the*

boy). Most of what we say concerning the production results for the alternators will concern *empty* and *stuff*, and not *splash*.)

Mean preference scores are listed for each alternator, as a function of query topic and age group, in Table 5. In order to assess whether or not a preference for one locative form over the other was significantly different from the null hypothesis of no preference, we performed a series of planned two-tailed *t*-tests on the difference of mean preference scores from zero, for each age group. The adults showed a significant preference for the *stuff*-content form over the *stuff*-container form ($M = 0.47$; $t(15) = 3.38$, $p < .005$), but they showed no significant preference for either form of *empty* or *splash*. Each of the child groups, as well, showed a significant preference for the *stuff*-content form ($M_{young} = 0.70$, $t(15) = 5.51$, $p < .001$; $M_{mid} = 0.59$, $t(15) = 4.54$, $p < .001$; $M_{old} = 0.50$, $t(15) = 3.16$, $p < .01$; $M_{combined} = 0.60$, $t(47) = 7.51$, $p < .001$). The combined group of children *deviated* from the adult standard, however, in their productions of *empty*: children significantly preferred the *empty*-content form over the *empty*-container form ($M = 0.23$, $t(47) = 2.06$, $p < .05$), despite the fact that adults showed absolutely *no* preference for either form ($M = 0$). Furthermore, the non-standard preference for the empty-content form was observed (though not significant) for each of the individual child groups. These (nonsignificant) preferences obtained for responses to the primary query as well as for responses collapsing across all three levels of query, with the mean preference score for primary responses approaching significance for the mid-aged children, ($M = 0.25$), $t(15) = 1.83$, $p \approx .09$. (See Table 4 for the frequency of *empty* locatives at different levels of response.)

On the issue of differences in mean preference score across age groups, we performed an Analysis of Variance on the mean preference score for each alternator, with the within-subject factor of Query Topic and the between-subjects factor of Age Group. For each verb, we found

Table 5

Mean Preference Score for Each Alternating Verb as a Function of Query Topic and Age Group

	AGE GROUP			
VERB	2;6-3;5	3;6-4;5	4;6-5;11	Adult
Empty				
Content-Topic Query	0.38	0.38	0.38	0.50
Container-Topic Query	0.06	0.19	0.00	-0.50
Mean	0.22	0.28	0.19	0.00
Stuff				
Content-Topic Query	0.78	0.81	0.81	0.81
Container-Topic Query	0.62	0.38	0.19	0.12
Mean	0.70	0.59	0.50	0.47
Splash				
Content-Topic Query	0.28	0.00	0.50	0.66
Container-Topic Query	0.22	-0.12	-0.47	-0.47
Mean	0.25	-0.06	0.02	0.09

Note: Mean preference score was calculated by subtracting the mean proportion of trials in which container locatives were produced from the mean proportion of trials in which content locatives were produced.

a highly significant main effect of query topic, showing a relatively greater preference for content locatives in response to content-topic queries than in response to container-topic queries: for *empty*, $M_{content\text{-}topic} = 0.41$, $M_{container\text{-}topic} = -0.06$, $F(1, 60) = 10.87$, $p < .005$; for *stuff*, $M_{content\text{-}topic} = 0.80$, $M_{container\text{-}topic} = 0.33$, $F(1, 60) = 21.81$, $p < .001$; for *splash*, $M_{content\text{-}topic} = 0.36$, $M_{container\text{-}topic} = -0.21$, $F(1, 60) = 29.62$, $p < .001$. These main effects demonstrate the efficacy of our method of using focused queries in order to elicit the full range of locative constructions. In addition, we also found an interaction of age group and query for the production of *splash* locatives, indicating that the difference in mean preference score in response to content- and container-topic queries was greater for subjects of increasing age, $F(3, 60) = 7.00$, $p < .001$. Follow-up *t*-tests revealed a significant difference in mean preference scores for the oldest children (mean difference = 0.97, $F(1, 15) = 22.01$, $p < .001$) and adults (mean difference = 1.13, $F(1, 15) = 22.09$, $p < .001$), but not for the younger children. This interaction suggests, quite plausibly, that the oldest children and adults are more sensitive than younger subjects to the discourse function of locatives.

What we *didn't* find were any significant main effects of age group. Furthermore, a series of planned two-tailed *t*-tests on the difference in mean preference score between child and adult groups also revealed no significant differences. In particular, we found no significant differences, between children and adults, in the preference for *stuff*-content over *stuff*-container forms (though children in every age group showed a greater preference than adults for the *stuff*-content over the *stuff*-container form); and no significant differences, between children and adults, in the preference for *empty*-content over *empty*-container forms (though, again, children showed a significant preference for the *empty*-content over the *empty*-container form).

In summary, we found that children in each age group were willing to produce a significant proportion of non-standard *fill*-content utterances, but were not willing to produce non-standard forms in general: *pour*-container and *dump*-container forms were not produced in significant proportions. Children also showed a preference for the production of *empty*-content over *empty*-container forms, despite the fact that adults showed no such preference. These results provide strong support for the claim that children prefer content locatives over container locatives, and thus replicate Bowerman's (1982) findings from spontaneous speech.

Comprehension. Turning to the results of the comprehension task, we will again view the adult performance as the standard, and subsequently focus on the occurrence or non-occurrence of *incorrect* locative interpretations. Crucial to our discussion will be the distinction between the two types of picture sets used in the study. The picture sets used for *pour/fill* and *dump/empty* were designed to assess a *bias* in the interpretation of the verb's meaning; that is, whether subjects preferred the manner or the endstate interpretation of the predicate. The picture sets for *splash* and *stuff*, on the other hand, don't provide us with a direct test of bias--just whether or not subjects were sensitive to a particular manner or endstate in the interpretation of a verb's meaning. Putting aside the issue of production for particular verbs, both types of comprehension tasks are potentially relevant, a priori, to the learnability question of what licenses unheard forms. In other words, if the range of possible forms that a child assigns to a verb is a function of the verb's meaning, then it may be a function of the *preferred* interpretations of the verb or a function of the *possible* interpretations of the verb. In either case, we will focus on differences in sensitivity or bias between the adult and child groups.

Let's begin with the bias tests for *pour*, *fill*, *dump*, and *empty*. In order to assess bias towards the manner or endstate interpretations of particular verb meanings, we set the following criterion: if a subject chose the same type of panel (manner or endstate) on both

comprehension trials for a given verb, then he or she was considered biased towards (and thus sensitive to) that interpretation of the verb's meaning. In Table 6, we've tallied the subjects, per age group, who were biased towards the manner or endstate interpretations of verb meaning. The underlined numerals indicate that the obtained frequency of biased subjects is significantly different from chance, at $p < .05$, according to a two-tailed binomial test. Just by chance, we'd expect a quarter of the subjects in each group to meet the criterion.

As is clear from Table 6, we found markedly different performance for *pour* and *dump* versus *fill* and *empty*. As we expected, adults unanimously treated *pour* and *dump* as having more to do with manner than endstate, and *fill* and *empty* as having more to do with endstate than manner. Furthermore, the number of children who were biased towards the standard interpretations of *pour* and *dump* was significantly higher than chance for every age group ($p < .05$). Yet this was not true of *fill* and *empty*. On the contrary, we found that none of the child groups were significantly biased towards the change of state interpretation of *fill* and *empty*, and furthermore that some of the child groups were biased towards the *incorrect* meaning of *fill* and *empty*. Specifically, eight of the oldest children (out of 16) were consistent in their interpretation of *fill* as having more to do with a pouring manner than a full endstate ($p < .05$). In addition, in every child group more children than would be expected by chance (though not significantly so) were sensitive to the incorrect meaning of *empty*; altogether, in fact, 19 children (out of 48) judged *empty* as having more to do with a dumping manner than an empty endstate, a significant result, at $p < .05$, in a two-tailed binomial test.

Although we didn't test subjects' sensitivity to different manners of filling (or emptying) in this experiment (see Experiment 2), it seems plausible to suggest that the willingness that children have in choosing a picture of pouring-spilling over a picture of dripping-filling, as the better instance of *filling*, derives in large part from their *particular* interpretation of *filling* as

Table 6

Frequency of Subjects Biased Towards the Manner or Endstate Interpretations of Verb Meaning

	AGE GROUP				
	2;6-3;5	3;6-4;5	4;6-5;11	Combined Children	Adult
MANNER BIAS					
Pour	13	14	13	40	15
Fill	2	4	8	14	0
Dump	9	15	16	40	15
Empty	6	7	6	19	0
ENDSTATE BIAS					
Pour	1	1	0	2	1
Fill	5	7	5	17	14
Dump	1	0	0	1	0
Empty	5	7	5	17	14

Note: A subject was counted as biased towards an interpretation if he or she chose the same type of panel (manner or endstate) on both comprehension trials for a given verb. Underlined numerals indicate that the obtained frequency of biased subjects is significantly different from chance, at $p < .05$, according to a two-tailed binomial test.

involving a pouring manner. In other words, many children may take the meaning of *fill* to include information about a pouring manner because the events of filling to which they are exposed are often events of pouring as well; pouring is a common means to the end of filling. Similarly, dumping is a common means to the end of emptying. Adults, by contrast, have sorted out that the causal connection between pouring and filling, or between dumping and emptying, is a fact about the world, and not a fact about the semantics and syntax of *fill*, or *empty*.

A further question, given these results, is the following: can we conclude that children are *generally* biased towards a manner interpretation, and against an endstate interpretation, of verb meaning? Of course, the generality of this claim demands the consideration of a larger sample of lexical items. Nonetheless, we performed a series of planned two-tailed *t*-tests, for different child groups, on the mean difference between the number of manner and endstate responses in trials for *pour*, *fill*, *dump*, and *empty*. Because this set of verbs is balanced with respect to the adult interpretation of verb meaning (the mean difference between manner and endstate responses for adults was 0.12, not significantly different from zero), we predict the following: if children are not biased in their interpretation of verb meaning, the difference between the proportions of manner and endstate responses should not be significantly different from zero. We found that children, in general, were highly biased towards a manner interpretation of verb meaning ($M_{combined} = 0.40$, $t(47) = 7.32$, $p < .001$). Furthermore, we found the same result for the children of each age group: for the youngest children, $M = 0.28$, $t(15) = 3.20$, $p < .01$; for the mid-aged children, $M = 0.39$, $t(15) = 4.28$, $p < .001$; for the oldest children, $M = 0.52$, $t(15) = 5.26$, $p < .001$. Children were more biased with increasing age, and there was a trend towards significance, at $p \approx .09$, in the difference in bias between the youngest and oldest children, $t(30) = -1.78$.

Although we have considered only four lexical items, our findings are a clear extension of Gentner's (1978) results with verbs like *mix*, *stir*, and *shake* (i.e., children had more difficulty in learning the meaning of *mix*, which specifies that a homogenous combination of substances be the result of an action, than in learning the meanings of *shake* or *stir*, which specify the particular motions involved in an action). We would like to suggest, therefore, that the particular misinterpretations of *fill* and *empty* (as involving pouring and dumping manners, respectively) are neither arbitrary nor isolated, but rather are due to a general bias towards interpreting verb meaning in terms of the manner in which a substance changes location. Notice that the postulation of a general manner bias accounts not only for why children make semantic errors with *fill* and *empty* (versus *pour* and *dump*) but also--in conjunction with the linking rule--for why children prefer to overgenerate content locatives. The question remains, however, as to what would motivate the child to search for properties concerning the manner in which substances change location. Part of our answer, following Gentner (1978, 1982), is that children are generally biased towards interpretations of verb meaning involving changes of location, versus changes of state, simply because the former are more perceptually salient. In the General Discussion, we shall take a closer look at the *particular* manners of locative verbs that are subject to this general bias.

In examining the results of the comprehension task for *splash* and *stuff*, we set the following criterion: if a child chose the standard panel on both comprehension trials for a given verb reading (manner or endstate), then he or she was considered sensitive to that reading. The standard choice, as usual, was defined by adult performance, which agreed with our own intuitions: adults were unanimously sensitive to the manner of splashing ('splashing with hands' was chosen over 'pushing with hands'), the endstate of splashing (a wet goal was chosen over a dry goal), the manner of stuffing ('forcing in' the contents was chosen over 'scooping up' the contents), and the endstate of stuffing (a bulging container was chosen over

a half-full container). In Table 7, we've tallied the number of subjects, per age group, who were sensitive to the standard or non-standard interpretations of verb readings. The underlined numerals indicate that the obtained frequency of sensitive subjects is significantly different from chance, at $p < .05$, according to a two-tailed binomial test. By chance, we'd expect a quarter of the subjects in each group to meet the criterion.

We are most interested in those cases where children are insensitive to the standard interpretation of a verb reading. What we find is that the youngest children are insensitive to the endstate and especially manner interpretation of *stuff*, and that the youngest and mid-aged children are insensitive to the endstate interpretation of *splash*. Table 7 also shows that none of the subjects were sensitive to non-standard manners or endstates, a result which is not surprising given the arbitrary choice of non-standard panels. (Compare the would-be manner of *splash* which involves pushing with the characteristic pouring manner of *fill*.)

The results for *splash* probably reflect the particular biases that English speakers have concerning the verb's meaning. Although we did not test for bias in the interpretation *splash*, on our semantic analysis the manner of splashing (i.e., involving the motion of an array of drops or particles) is more important or constrained in the meaning of the verb than is its endstate; for example, in uttering *John splashes water at Mary*, a speaker implies that Mary may not have become wet as the result of John's splashing. By contrast, it is not possible, in using the verb *splash*, to imply that the content may not have changed location in a characteristic manner. A further possible cause for the insensitivity of the younger children to the endstate reading of *splash* is that young children, in general, may have difficulty identifying the changes of state selected by predicates.

In the case of *stuff*, it is also possible that the insensitivity of the youngest children to the

Table 7

Frequency of Subjects Sensitive To the Standard or Non-Standard Interpretations of Verb Readings

	AGE GROUP				
	2;6-3;5	3;6-4;5	4;6-5;11	Combined Children	Adult
STANDARD SENSITIVITY					
Stuff$_{mnr}$	5	8	10	23	16
Stuff$_{end}$	7	10	14	31	16
Splash$_{mnr}$	13	14	16	43	16
Splash$_{end}$	7	6	13	26	16
NON-STANDARD SENSITIVITY					
Stuff$_{mnr}$	4	4	1	9	0
Stuff$_{end}$	2	2	0	4	0
Splash$_{mnr}$	0	0	0	0	0
Splash$_{end}$	1	4	0	5	0

Note: A subject was counted as sensitive to an interpretation if he or she chose the same type of panel (standard or non-standard) on both comprehension trials for a given verb reading. Subscripts on verbs indicate verb readings. Underlined numerals indicate that the obtained frequency of biased subjects is significantly different from chance, at $p < .05$, according to a two-tailed binomial test.

endstate reading of the verb was due to a general insensitivity to the endstate interpretations of verb meaning. The insensitivity to the youngest children to the standard manner of stuffing, however, is more difficult to explain. Unlike the case with *splash*, both the manner and endstate of stuffing appear to be mutually constraining: for example, a speaker who says *John stuffed the box with toys* is probably implying that the toys themselves were handled in a fashion characteristic of stuffing: they were *forced* into the box, not dropped or thrown. Conversely, the manner reading seems to imply the endstate reading. That is, why would a stuffing manner be employed unless the capacity of the container was exhausted? Of course, the crucial point is whether or not these intuitions are shared by every speaker who knows the meaning of the verb *stuff*. Our best guess is that they are. The question, then, remains: why are young children insensitive to the standard manner of *stuff*. One answer is that the result is an artifact of the panels that we used in testing the manner of stuffing: the youngest children had particular trouble in understanding the non-standard manner panel which showed an agent 'scooping up' the contents, using the container both as an instrument and as a goal; many gave clear evidence (i.e., a spontaneous description) of interpreting such panels as depicting an action in which the container was treated as a source instead of a goal, and in which the contents were being "dumped" or "thrown out" instead of stuffed into the container (as if the "motion lines" of the drawing were misread). Of the sixteen youngest children, eight misinterpreted at least one of the two 'scooping-up' panels in this way, and four misinterpreted both panels in this way. Curiously, of the total 12 trials in which the youngest children misinterpreted this panel, 10 *chose* that panel in the comprehension task. Fortunately, responses in which the drawings themselves were overtly misunderstood were rare (25 out of 768 child trials across age groups, or 3.3%), with most of them (18/25 or 72%) involving the 'scooping-up' panels.

In summary, we found that children are prone to harboring particular misinterpretations of *fill* and *empty*, according to which typical manners of filling and emptying--the pouring of a substance into a container and the dumping of a substance out of a container, respectively--are judged to be more important in the meanings of the verbs than the endstates of the container. We argued that these particular misinterpretations are a reflection of a general bias to interpret the meanings of verbs in terms of changes of location, rather than changes of state. In support of this conclusion, we found that while some of the child groups were biased towards the manner interpretations of *fill* and *empty*, none of the child groups were biased away from the manner interpretations, or towards the endstate interpretations, of *pour* and *dump*. Further support for this treatment of verb misinterpretations comes from our results with *splash*, in which we found that younger children were selectively insensitive to the endstate reading of *splash*, and suggested that the endstate interpretation is less constrained according to adult intuitions and perhaps less accessible to young children on general grounds.

Association. If it is indeed the case, as hypothesized, that verb meaning and syntax are linked in the lexicons of language learners, we would expect to find a correspondence between syntactic and semantic errors with locative verbs. Crucially, given that these are real verbs, showing just *any* association between syntax and semantics is trivial--after all, because there is a correspondence between syntax and semantics in adult speech, we would expect a similar correspondence in child speech *just on the basis of input.* The question, then, is: are syntax and semantics associated in cases of novel (non-standard) usage?

At this point, we can already provide some unequivocal support for the linking hypothesis simply by noting that those verbs which were subject to the children's preference for content locatives (*fill* and *empty*) are the same verbs which were subject to the children's bias towards manner interpretation; by contrast, *pour* and *dump* were rarely the sources of syntactic or

semantic errors. Regardless of what our specific tests of association may yield, we argue that this general association of syntactically overgenerated and semantically overextended or misapplied predicates must be explained by any theory of verb acquisition.

The results from the production task, and the design of the comprehension task, leave us with the following specific tests of association: for *fill*, we can assess the occurrence of non-standard forms as a function of a bias in interpretation; for *empty*, we can assess the non-standard preference for a locative form as a function of a bias in interpretation. Other specific tests of association, although permitted by our design, have been ruled out by the fact that children failed to utter non-standard forms: due to the production results for *pour* and *dump*, we cannot assess the occurrence of non-standard forms as a function of a bias in interpretation; similarly, for *stuff* and *splash*, we cannot assess the non-standard preference for a locative form as a function of insensitivity to an interpretation.

A 2 x 2 contingency table was constructed for *fill*, with each child (per age group) contributing one score to the table. The semantic levels of the table were defined as follows: a child was scored as biased towards manner if both responses in the forced-choice task identified *filling* more with a pouring manner than a full endstate; a child was scored as biased towards endstate if both responses in the forced-choice task identified *filling* more with a full endstate than a pouring manner. For the syntactic levels, a child was scored as either producing at least one *fill*-content form, or producing no *fill*-content forms. (In other words, the production of *any* fill-content forms was taken to be critical; a numerical preference for *fill*-content forms over *fill*-container forms by the children was judged too strong a test, given the positive input of *fill*-container forms that children undoubtedly receive from parents. In addition, two children who produced *other* forms in the *fill* production trails were eliminated from the analysis.) We found, for the oldest children (4;6-5;11), that there was a trend ($p <$

.08) towards association according to a Chi-Square test ($X^2(1, N = 12) = 3.09$): a child who is biased towards the manner interpretation of *fill* will tend to produce its content locative; on the other hand, a child who is biased towards the endstate interpretation of *fill* will not tend to produce its content locative. We note, however, that this table of counts failed to reach significance by a Fisher Exact Test (one-tailed $p \approx .12$). The contingency table for each child group and for the combined child groups is presented in Table 8. No significant associations were found for the other child groups or for the combined child groups.

To a first approximation, then, for the group of children who were most biased towards the manner interpretation of verb meaning (according to our comprehension results), we found that instances of overgeneration were somewhat associated with corresponding misinterpretations of verb meaning. This conclusion is tentative, however, in that the observed association is only probabilistic: of the 12 oldest children who were counted in the contingency table, three appear to have violated the predicted linking of syntax and semantics. Across age groups, in fact, sixteen children (out of 29) appear to have violated the hypothesized linking (11 were biased towards endstate yet produced at least one *fill*-content form; 5 were biased towards manner yet produced no *fill*-content forms.) Clearly, if children do indeed make use of linking rules in learning the syntax of new verbs, we have only managed a very crude test of these rules. At this point, let us make three observations which will put these results--especially the unpredicted findings--in perspective. First, some of the noise in the data is undoubtedly due to the fact that some children have misinterpreted *fill* in a way other than we predicted. In other words, given that children are biased towards a manner interpretation of *fill*, there is no guarantee that we picked the *right* manner interpretation. For reasons outlined in the introduction, we have simply settled on the 'pouring' interpretation of *fill*. A different manner interpretation of *fill*, however, is that the top surface of a substance moves higher and higher during the course of the action. Notice that a child who is biased towards this *content-up*

Table 8
Contingency Tables per Age Group on the Association of Bias in the Interpretation of *Fill* and the Occurrence of *Fill*-Content Locatives

2;6-3;5:

	Produced At Least One Content Locative	Produced No Content Locatives
Manner Biased	1	1
Endstate Biased	4	1

3;6-4;5:

	Produced At Least One Content Locative	Produced No Content Locatives
Manner Biased	2	2
Endstate Biased	6	0

4;6-5;11:

	Produced At Least One Content Locative	Produced No Content Locatives
Manner Biased	5	2
Endstate Biased	1	4

COMBINED CHILDREN:

	Produced At Least One Content Locative	Produced No Content Locatives
Manner Biased	8	5
Endstate Biased	11	5

Note: Each child counted in a table contributed one score to that table. A subject was scored as biased towards an interpretation if he or she chose the same type of panel (manner or endstate) on both comprehension trials for *fill*.

interpretation of *fill* would choose the full endstate panel in the comprehension task--and would thus be indistinguishable from a child who was genuinely biased towards endstate. Moreover, some evidence for this misinterpretation of *fill* comes from the frequency with which children used the particle *up* in content locatives--as we've seen, children used *up* equally often in content and container locatives, despite the fact that adults only use the particle *up* (versus the preposition *up*; see Talmy, 1985) to emphasize the completeness of a change of state (e.g., *John loaded up the truck with the hay*; ?*John loaded up the hay into the truck*). Children who say things like *fill the water up*, we suggest, are not only using *up* literally in such forms, but also taking *fill* to be a change of location verb. Thus, of the eleven children who produced at least one *fill*-content form, but were not biased towards the pouring misinterpretation, some could have had another misinterpretation of *fill* in mind. In fact, of these 11 children, three produced at least one content locative containing the particle *up*. (Note furthermore that the actual utterance of *up* in *fill*-content forms presumably underestimates the frequency of the content-up interpretation.)

Another portion of these eleven children can be accounted for in the following way: a child who is biased towards the endstate interpretation of *fill* may still be *sensitive* to the pouring interpretation of *fill*--a possibility that we did not test. Crucially, if it is sensitivity to, rather than bias towards, a manner interpretation of verb meaning which licenses the content locative, our test of association would be inappropriate (i.e., in this comprehension task, manner sensitivity is underestimated by manner bias; manner insensitivity is overestimated by endstate bias). One argument in favor of regarding sensitivity to an (essential) component of meaning as the relevant criterion is that bias towards one reading or the other of alternators does not seem to rule out alternative locative forms. For English locative verbs which alternate (e.g., *inject*, *smear*, *spatter*, *spread*, *sprinkle*, *stack*, *cram*, *crowd*, *jam*, *shower*, *wrap*, *load*), there is often

the intuition that one reading or the other is more basic. Thus, we argued above that adults are probably biased towards the manner interpretation of *splash*. Similarly, on our semantic analysis, adults who use the verbs (e.g.) *inject*, *smear*, *spread*, and *sprinkle*, may intend to communicate more about the content's change of location than about the container's change of state; conversely, perhaps, in the cases of (e.g.) *wrapping* and *loading*. Yet, for all of these presumed biases, these verbs are alternators nonetheless. A second argument in favor of retreating from the stronger criterion of bias to the weaker criterion of sensitivity is simply that bias assumes the competition of two interpretations--an assumption which would leave the child unable to use linking regularities in those cases where only one interpretation (e.g., change of location in a certain manner) is thought by the child to be relevant to the meaning of a verb (e.g., *pour*). We will return to this issue in Experiment 2, where we will argue that *both* sensitivity and bias tests are needed to isolate components that are essential to the meaning of a verb.

A third observation is simply that as children grow older, they hear more and more utterance of *fill*-container forms, which would selectively discourage older children from uttering *fill*-content forms despite any influence of linking. Of course, that English learners ultimately *stop* producing *fill*-content forms does not follow from the absence of *fill*-content forms in parental speech, given the lack of negative evidence. We do assume, however, that children are sensitive to different locative forms--in the sense that the utterance of *fill*-container forms becomes more routine than the utterance of *fill*-content forms. It is probable, therefore, that of the five children who were sensitive to manner, but failed to produce at least one content locative, some may have suffered from the influence of positive input (i.e., we can't force a child to give us a *fill*-content form, even if it is within his or her capacity).

Before turning to the association test for *empty*, we note that the semantic errors for *fill* (which reach their peak of 59% for the oldest group of children) lag behind the syntactic errors (the oldest children produced fewer *fill*-content forms (11 utterances, $M_{old} = 0.34$) than did the younger children (17 utterances for each of the younger groups, $M = 0.53$)). Although the difference in *fill*-content production between child groups is not significant, this lag between the types of errors raises the question of whether syntax can license semantics rather than the other way around. One possible scenario, consistent with this hypothesis, is the following: younger children would overgeneralize the content locative to *fill* because the content locative is the dominant locative pattern in English; older children, on the basis of the overgeneralized *fill*-content form, would then search for an interpretation of *fill* that involves a change in the location of the content--in effect, satisfying the linking regularities in reverse. Although this explanation is consistent with our findings, a simpler account of the gap would involve just the hypothesis of the linking regularity--used in the hypothesized direction, from semantics to syntax--plus the observation that verb learning (or testing) doesn't occur in a vacuum: as children grow older, they not only receive more positive input (which would selectively discourage older children from uttering *fill*-content forms despite any influence of linking, as discussed above), but also become more determinate in their comprehension responses (the number of subjects who "split" their comprehension responses drops from nine (youngest), to five (mid), to three (oldest)). In combination, these factors of positive input and growing decisiveness would result in the artifact that the bulk of syntactic overgeneration in our production task precedes the peak of semantic misinterpretation in our comprehension task. In summary, therefore, the results do not bear directly on whether semantics licenses syntax or vice versa.

On the association test for *empty*, we again constructed a 2 x 2 contingency table, with each child (per age group) contributing one score to the table. The semantic levels of the table were

defined as in the tables for *fill*: a child was scored as biased towards manner if both responses in the forced-choice task identified *emptying* more with a dumping manner than an empty endstate; a child was scored as biased towards endstate if both responses in the forced-choice task identified *emptying* more with an empty endstate than a dumping manner. For the syntactic levels, however, we departed from the method used above; instead of taking the production of *any* content locatives to be critical, we scored each child on whether he produced more (i.e., only) *empty*-content forms or more (only) *empty*-container forms. These syntactic categories are appropriate given the finding that adults in our task produced equal numbers of either locative form, and the assumption--on this basis--that parental input also shows no preference for either locative form. In fact, given our own intuitions that the *empty*-container form is more frequently produced by adults in spontaneous speech we suggest that the association test is a conservative one.

We found a significant association, for the combined group of children, according to a Fisher Exact test (one-tailed $p < .02$): a child who is biased towards the manner interpretation of *empty* will tend to produce more *empty*-content locatives than *empty*-container locatives; on the other hand, a child who is biased towards the endstate interpretation of *empty* will tend to produce more *empty*-container locatives than *empty*-content locatives. In addition, we found a trend towards significance for the mid-aged children (3;6-4;5) according to a Fisher Exact test (one-tailed $p \approx .07$). The contingency table for each child group and for the combined child groups is presented in Table 9. No significant associations were found for the other child groups.

These results provide support for the claim that verb meaning and verb syntax are linked in some fashion; of the 15 children who were biased towards manner (in the combined table), 14 produced only *empty*-content forms. By contrast, for those children who were biased towards

Table 9
Contingency Tables per Age Group on the Association of Bias in the Interpretation of *Empty* and the Preference for *Empty*-Content or Empty-Container Locatives

2;6-3;5:

	Produced Only Content Locatives	Produced Only Container Locatives
Manner Biased	3	1
Endstate Biased	2	1

3;6-4;5:

	Produced Only Content Locatives	Produced Only Container Locatives
Manner Biased	5	0
Endstate Biased	3	4

4;6-5;11:

	Produced Only Content Locatives	Produced Only Container Locatives
Manner Biased	6	0
Endstate Biased	0	1

COMBINED CHILDREN:

	Produced Only Content Locatives	Produced Only Container Locatives
Manner Biased	14	1
Endstate Biased	5	6

Note: Each child counted in a table contributed one score to that table. A subject was scored as biased towards an interpretation if he or she chose the same type of panel (manner or endstate) on both comprehension trials for *empty*.

endstate, there was no overall preference for one locative form or the other (i.e., of the 11 children who were biased towards endstate, 5 produced only *empty*-content forms and 6 produced only *empty*-container forms). These results suggest that it is a bias towards manner, but not endstate, which is linked to syntax; if a child is biased towards a specific manner (dumping) in the interpretation of *empty*, the argument corresponding to the content must be encoded as the direct object (and to some extent, the converse). As with the results of the association test for *fill*, however, we argue that sensitivity, rather than bias, may be the more appropriate criterion. In particular, of the five children who were biased towards the endstate interpretation of *empty*, but who produced only content locatives, some may have been sensitive to the dumping (or some other) manner of *empty*. More generally, the hypothesis of linking between verb meaning and verb syntax must show, for every verb capable of alternation, that two readings exist for the adult speaker. In some cases, such a demonstration is hampered by the obvious bias towards one reading or the other; in the case of *empty*, we would argue, adults prefer the change of state interpretation. Such bias, however, does not rule out sensitivity to an interpretation involving the manner in which a substance changes location. (For adults, the interpretation of *empty* which is relevant to the content may be more abstract than a change in location; it may have something to do with the fact that when one empties the garbage, let's say, one is *disposing* of the garbage, changing its availability.)

In summary, our tests of association have provided support for the linking of verb meaning and verb syntax in the lexicons of first language learners: for the oldest children, we found that instances of overgeneration of *fill* to the content locative form were somewhat associated with the misinterpretation of its meaning as having more to do with a pouring manner than a full endstate. We argued that the strength of association was weakened by the insensitivity of our comprehension task to other manners (besides pouring) of filling, by the design of our

comprehension task to assess bias rather than sensitivity, and by the influence of positive input of the container locative form. For the combined group of children, we found that the non-standard preference for *empty*-content forms over *empty*-container forms was strongly associated with the misinterpretation of its meaning as having more to do with a dumping manner than an empty endstate. Here again, we considered the criterion of bias to be possibly too strong, and suggested instead that the criterion of sensitivity was more appropriate.

We introduced this experiment with three questions: will children make syntactic errors? Will they make semantic errors? And will the syntactic and semantic errors be associated with one another? As we have shown, the answer to each question is *yes*: children overgenerate the content locative; they are biased towards the manner interpretation of verb meaning; and the instances of syntactic overgeneration and semantic misinterpretation are associated with each other. Although we don't have direct evidence on how this linking is used, these answers cast strong doubt on the sufficiency of any purely syntactic account of how verb errors arise (e.g., the overregularization of the NP-V-NP-*into/onto*-NP form). On the contrary, if syntactic accounts are too simple, but must be supplemented by the linking of syntactic form and verb meaning, why can't this linking be used, in principle, in the service of verb learning? In the second experiment, we attempt to replicate, and indeed strengthen, the association between syntactic and semantic errors with *fill* by testing for children's sensitivity to the pouring manner of filling.

Experiment 2

In our second experiment, we attempted a "case study" of one verb -- *fill* -- a study in which we conducted sensitivity as well as bias tests of comprehension in order to show a stronger association of elicited syntactic and semantic errors. Such a concentrated study of one lexical item has obvious advantages and disadvantages. On the one hand, we avoid any systematic confusion between verbs; children are better able to focus their attention on one verb than on six. On the other hand, we must take care not to overgeneralize on the basis of the idiosyncratic properties of *fill*. In this regard, we hope to show (in the General Discussion) that all of the available evidence suggests the generality of linking, not only as the source of--and ultimately, solution to--the mistakes children make with locative verbs, but also as an essential feature of languages across the world.

Besides focusing on the production and comprehension of just one verb, we made three other major changes in this experiment from the last. First, we dispensed with the practice of initially presenting a drawing, labelled as (e.g.) filling, before each forced choice; this procedure was no longer needed to focus the child's attention on the current verb, and by omitting it we also eliminated any potential for introducing a response bias towards the "original" manners and endstates. Second, we used three-panel pictures instead of two-panel pictures. This modification insured, we thought, that any observed bias would have more to do with the "message" than the "medium"; that is, it is less likely that the format of the drawings would lead children to ignore the endstate panel simply because they don't understand its relation to the action being depicted. Third, the production and comprehension tasks were completely separated, with the block of production trials following the block of comprehension trials. Furthermore, the production task no longer consisted of describing, albeit in a structured way, the same materials used in the comprehension task. This complete separation of tasks

permitted us to develop and use a more powerful production technique, and one less influenced by the semantic and pragmatic focusing of the comprehension task.

Method

Subjects. Forty-eight children and sixteen adults, all native speakers of English living in the Boston area, participated in the study. As in the previous study, the children fell into three age groups; however, in the current experiment we decided to test a broader range of children, with more discrete age groupings. Our rationale was that the condensed age range of the first experiment may have been too fine-grained to snow any clear developmental differences and too early to show when children acquire adult competence. Therefore, we tested sixteen children between the ages of 3;5 and 4;6 (mean 4;0); sixteen between the ages of 4;9 and 6;6 (mean 5;7); and sixteen between the ages of 6;10 and 8;9 (mean 7;9). (Four children were replaced in the design for failing to attend to the comprehension task.) The children were drawn from middle-class day-care and after-school programs in Cambridge, Needham, and Newton. The adult subjects were undergraduates at MIT, ranging from roughly 18 to 22 years of age, and were paid for their participation.

Materials. For the comprehension task, each subject was shown twelve pairs of drawings; for the production task, each subject was shown four actions, each involving one of four contents and one of four containers.

Each drawing used in this experiment was composed of three panels, where the first panel depicted a beginning or early state of the action, the second panel depicted a mid state of the action, and the third panel depicted the endstate of the action. In a drawing of 'pouring -&- spilling' (Figure 7), for example, the first panel shows a woman in the process of pouring water from a pitcher, but with a small puddle appearing next to an empty glass; the second

panel shows the woman continuing to pour water from the pitcher, but with a larger puddle next to the empty glass; the third panel shows that the woman has left the scene (the action presumably over), leaving behind an even larger puddle next to the empty glass. To take another example, in a drawing of 'dripping -&- filling' (Figure 8), the first panel shows a woman in the process of turning on a faucet, allowing water to drip from a spigot into a half-filled glass; the second panel shows the woman continuing to allow water to drip from the faucet into the glass, now three-quarters filled; the third panel shows that the woman has left the scene (the action over), leaving behind a glass full of water. The reader may want to contrast these drawings with the comparable two-panel drawings used in Experiment 1 (Figures 1 and 2); we believe that the new three-panel drawings are easier to interpret, any commentary aside. (Whereas in Experiment 1 the first panel depicted a state of the action showing manner and the second panel depicted endstate, in this experiment the depiction of manner and endstate is more "dynamic" by virtue of being distributed across more than one panel: the manner is represented in the first two panels and the accomplishment of the endstate is represented in all three panels.)

The twelve pairs of drawings shown to each subject were designed to test three distinct types of judgments that subjects can make about the meaning of *fill*: sensitivity to particular manners of filling; sensitivity to particular endstates of filling; and bias towards particular manners versus particular endstates of filling. Of the twelve pairs of drawings shown to each subject, two tested sensitivity to manner, six tested sensitivity to endstate, and four tested bias. The manners and endstates contrasted in these tests are listed in Table 10.

Sensitivity tests work either by holding constant the endstate of filling and contrasting two manners (manner sensitivity tests), or by holding constant the manner of filling and contrasting two endstates (endstate sensitivity tests). Specifically, the manners contrasted in tests of

Figure 7

Pouring -&- Spilling

Figure 8

Dripping -&- Filling

Table 10

Manners and Endstates Contrasted in the Comprehension Tests

	DRAWING 1			DRAWING 2	
	MANNER	ENDSTATE	X	MANNER	ENDSTATE
TYPE OF TEST					
Manner Sensitivity					
(cs1)	pouring	full	X	dripping	full
(cs2)	pouring	empty	X	dripping	empty
Endstate Sensitivity					
(cs3)	pouring	full	X	pouring	3/4-full
(cs3)	pouring	3/4-full	X	pouring	empty
(cs3)	pouring	empty	X	pouring	full
(cs4)	dripping	full	X	dripping	3/4-full
(cs4)	dripping	3/4-full	X	dripping	empty
(cs4)	dripping	empty	X	dripping	full
Bias					
(cs1)	pouring	full	X	dripping	empty
(cs2)	pouring	full	X	dripping	empty
(cs1)	dripping	full	X	pouring	empty
(cs2)	dripping	full	X	pouring	empty

Note: each line corresponds to the presentation of two drawings in a forced choice trial of the comprehension task. The manner and endstate depicted in each drawing are listed on either side of the 'X'. The parenthetical information identifies the contrast set of each test. Nothing about the temporal order of the trials, or the positions of the choices within each trial, is implied by the layout of this table.

manner sensitivity were pouring and dripping. This contrast was presented twice to each subject--once with an empty endstate (i.e., the contents spilled) and once with a full endstate.

In designing tests of endstate sensitivity, we went beyond the simple distinctions of full and empty that were used in the last experiment. In this experiment, we attempted to distinguish between subjects who conceived of filling as an *accomplishment*--an action of duration with a definite endpoint--versus those who saw the action merely as an activity or process, unmarked by any definite change in the state of a container (Vendler, 1967; Tenny, 1988). Therefore, the endstates contrasted in tests of endstate sensitivity included not only an empty container (succeeding panels show puddles of increasing size) and a full container (succeeding panels show a 1/2-full, 3/4-full, and full container), but also a three-quarters full container (succeeding panels show a 1/4-full, 1/2-full, and 3/4-full container). A drawing with a 3/4-full container is presented in Figure 9. By controlling the amount of content explicitly transferred in the drawings of full and 3/4-full containers (the only difference being the amount of content initially in the container), we could partially tease apart the accomplishment and process interpretations of filling in the following way: if a subject views filling as an accomplishment, he or she will always choose the full container over the 3/4-full container (though not, perhaps, vice versa). Put another way, if a subject always chooses the full and 3/4-full containers over the empty containers, but does not consistently choose the full containers over the 3/4-full containers, he or she is probably insensitive to filling as an accomplishment. (Note that we colored in the contents and containers across the panels in these drawings, but not in the bias or manner sensitivity drawings, in order to accentuate their relation to one another; subjects could at least *perceive* the extent to which a container was filled.) Each endstate contrast (e.g., a full container vs. an empty container) was presented

Figure 9

Pouring -&- 3/4 Full

(Endstate Sensitivity Test)

once with a pouring manner and once with a dripping manner, yielding a total of six endstate trials per subject.

Bias tests work by contrasting both the manner and endstate of filling. The contrasted manners were pouring and dripping; the contrasted endstates were an empty and a full container. Subjects were therefore shown two contrasts: between 'pouring -&- spilling' and 'dripping -&- filling' (a replication of the bias test in Experiment 1); and between 'pouring -&- filling' and 'dripping -&-spilling'. Within subject, each of these contrasts was presented twice, accounting for a total of four bias trials.

As in Experiment 1, we presented subjects with drawings of different "scenarios" in order to control for the salience of contents and containers in particular drawings. In fact, the same four scenarios (defined by agent, potential content, and potential container) used in Experiment 1 for the testing of *pour* and *fill* were used here: the panels comprising scenarios A1, A2, B1, and B2 (see Table 2) were modified for their current use--by varying the amount and location of the content (and the use of color for the endstate drawings). Thus, the bias contrast between 'pouring -&- spilling' and 'dripping -&- filling' could be presented via Figures 7 and 8, involving a woman, water, and a glass, or in one of three other scenarios: a man potentially filling a sink with water; a girl potentially filling a bowl with honey; and a boy potentially filling a bucket with paint. In order to balance the combination of scenarios and meaning contrasts, we divided the set of twelve contrasts into four subsets of three contrasts: the manner contrast, with a full container as the endstate, and the two bias contrasts; the manner contrast, with an empty container as the endstate, and the two (repeated) bias contrasts; the three endstate contrasts with a pouring manner; the three endstate contrasts with a dripping manner. This subdivision of the meaning tests insures that the two presentations of a given contrast (i.e., the two tests of a contrast which involves a change in the non-critical component of meaning or no

change at all) were made with different picture sets. We balanced the combination of picture set (scenario) and contrast set (meaning tests) across subjects within an age group.

For the production task, each subject was shown four actions of filling, each involving one of four contents and one of four containers. Due to the demands of the production procedure (described below), it was necessary to perform the actions with pairs of "inter-changeable" contents and containers. Furthermore, to enliven the production task and discourage rote responses, we used two subsets of materials: a "solid substance" set (consisting of marbles and pennies as contents; a bowl and jar as containers) and a "liquid" set (consisting of apple-juice colored water and grape-juice colored water as contents; a glass and cup as containers). The actions of filling performed for each subject rotated through all four contents and containers.

Procedure. The procedure consisted of first testing each subject's sensitivity to, and/or bias towards, interpretations of *fill*, and then eliciting from the subject locative forms of *fill*. The production task always followed the comprehension task because subjects were exposed to actions of filling during the production phase, actions which would have influenced any subsequent comprehension responses. Subjects were tested in a single session by two experimenters (one eliciting responses; the other observing), in an area as free as possible of potential distractions.

The sessions began immediately with the comprehension trials. (We omitted any separate introduction to the format of the pictures.) The comprehension trials consisted of twelve forced choices between pairs of drawings. For each forced choice, the experimenter began by placing the pair of drawings directly in front of the subject. The experimenter then "talked the subject through" each drawing, particularly the child subjects, before any forced choice was made. For the children, the experimenter began by having the child identify each of the "constituents" across the panels in a drawing. This procedure, in our judgment, was effective in helping

children see the continuity of constituents from one panel to the next. As an example of this procedure, the experimenter would say to a child: "Look at the first picture (Figure 7): point to the woman ...(experimenter waits for response); point to the pitcher ...; point to the water ...; point to the glass ... Now look at the second picture (Figure 8): point to the woman ...; point to the faucet ...; point to the water ...; point to the glass ..." If a child failed to point out an instance of a constituent, the experimenter would do so. After thus introducing the constituents in a panel, the experimenter would further reinforce the cohesion of the panels by saying (to adults as well as to children), "this is the beginning (experimenter points to first panel), the middle (experimenter points to second panel), and the end (experimenter points to third panel). When the woman does *this* (experimenter sweeps finger across first and second panels), it ends up like *that* (experimenter points to third panel)." Finally, after both drawings have been reviewed, the experimenter would ask, "Which of *these* (experimenter gesturing with both drawings) is *filling*?" If a subject did not clearly indicate either one drawing or the other, the experimenter repeated the question.

The order in which the twelve pairs of forced choices were presented was randomized separately for each subject within an age group. We ruled out randomized orders which resulted in two consecutive trials involving the same scenario, two consecutive trials testing the same contrast, or three consecutive trials testing the same *type* of contrast (manner sensitivity, endstate sensitivity, or bias). In addition, we also controlled for the position (left or right) of the alternative drawings in the forced choices; the position of choices was always balanced for the two presentations of a given contrast, so that each drawing in a contrast was presented once on the right and once on the left.

After completing the comprehension task, each subject participated in four trials on the production of locatives with *fill*. The production technique itself is like that used in Experiment

1 in that the primary queries used to elicit locatives make either the content or the container the topic of "discourse." Unlike these earlier production tasks, however, the current technique goes further in structuring the discourse and in making the *full* utterance of a content- or container-locative most felicitous. (We also collected twice as much production data in this experiment than in the last.) Much of the elaboration of the technique is due to the method of Crain, Thornton, and Murasugi (1987), used to elicit full passives (i.e., those with *by*-phrases) from three- and four-year olds. Essentially, Crain *et al* induced children to utter full passives by giving children a choice between two entities, only one of which participated in an action; children were then forced by discourse constraints to utter a *by*-phrase in order to identify the participant (e.g., in a context where one of two soldiers is bitten by an alligator, a child asked, "Which one is getting bited by the alligator?"). Similarly, we presented subjects (adults as well as children) with two potential contents or two potential containers in each production trial. Furthermore, the need for subjects to identify the actual participant in the action was impelled by their task of describing events of filling to a *blindfolded* puppet; the artifice of a blindfolded puppet forced the subjects to be explicit about the actions they were witnessing. These measures, supplemented by the additional discourse constraint that either a content- or container-topic query was posed, were designed to elicit full content- or container-locatives from subjects.

An example will clarify our production technique. At the close of the comprehension task, subjects were introduced to a blindfolded puppet and told by the experimenter, "I'm gonna do some filling, and I want you to tell Marty the Puppet what I'm doing." In each of the four succeeding trials, the experimenter would first identify the *topic* of the trial, then present the subject with two potential non-topicalized participants (i.e., two containers if the content was topicalized; two contents if the container was topicalized), then perform the action of filling with the topic and one of the two other participants, and finally pose the query. A typical trial

might be presented as follows: "Here are some marbles (topic). I can have either a jar or a bowl. Now watch this: I'm filling (experimenter fills jar with marbles) ...Can you tell Marty what I did to the *marbles*?" Notice that a pragmatically natural response in this context (standard syntax aside) is a content locative form, (e.g.) "You filled the marbles into the *jar*," where the old information (topic) has been encoded as the direct object and the new information--the chosen participant--has been encoded as the oblique object. Another example, in which a container is topicalized, is the following: "Here is a glass (topic). I can have either grape juice or apple juice. Now watch this: I'm filling (experimenter fills the glass with apple juice) ...Can you tell Marty what I did to the *glass*?" In this context, the most appropriate response is "You filled the glass with *apple juice*." If a subject failed to indicate an unambiguous direct object (or used a verb other than *fill*), the experimenter repeated the query, reminding the subject that Marty couldn't see (or telling the subject to "use the word *fill*"). As in the other experiments, we tested production at three levels of response, if necessary: in those trials where the primary query failed to elicit an unambiguous direct object, we followed up with the secondary prompt "filling (what)?"; in those trials where the secondary query failed to elicit an unambiguous direct object, we followed up with the tertiary prompt (e.g) "filling the glass or filling the apple juice?" (with order of adjuncts balanced within subject).

Within this structured production task, we made every effort to control for any effects due to the order of the trials or to the salience of the four contents and containers. Recall that two subsets of materials were used--a solid substance set (consisting of marbles and pennies as contents; a bowl and jar as containers) and a liquid set (consisting of apple-juice colored water and grape-juice colored water as contents; a glass and cup as containers). By strictly alternating between trials with one type of material or the other (within subject), we lessened the possibility of interference across trials. In fact, four orders of trials within the production task were employed across the subjects within an age group; the orders were determined by the

factorial combination of whether subjects began the production task with the solid set or the liquid set (alternating thereafter) and whether the order of query topics was content-container-container-content or container-content-content-container. Note that we counterbalanced the order of production trials with the combination of picture set and contrast set in the comprehension trials, on the grounds that the independence of the comprehension and production *tasks* is a prerequisite to testing the association of verb meaning and syntax.

Regarding the salience of particular contents or containers, we took the following precautions. First, we performed actions of filling for each subject with four non-overlapping pairs of contents and containers, two from the solid set and two from the liquid set, insuring that each of the four contents and containers was used once and only once per subject in the actions of filling. Second, the four possible pairings of solid contents and containers (counting whether the content or container was the topic) were counterbalanced with the four possible pairings of liquid contents and containers (again, counting whether the content or container was the topic), to yield 16 unique combinations of materials and query topics across the subjects within an age group. In addition, these pairings of content and container were counterbalanced with the order of production trials across subjects within an age group.

Scoring. Responses in the comprehension task were scored according to the manner and/or endstate of the chosen drawing in each forced choice trial, depending on the presented contrast. For the production trials, responses were scored according to whether the direct object of an acceptable *fill* locative form corresponded to the content or the container in the performed action. (Acceptable locative forms also included two passives (e.g., *the glass was filled up*, where *the glass* was scored as the true direct object) and two unaccusative intransitives (e.g., *the glass filled with marbles*, where *the glass* was scored as the true direct object; see the General Discussion for our assumptions about lexicosyntactic representation).) We relied on

the conventions and protocol listed in the Scoring Section of Experiment 1 for determining whether or not the use of a pronoun (e.g., *filling it*) was ambiguous. Responses which were undecipherable or not clearly locative (e.g., intransitive responses such as *she's filling with the juice*) were coded as *other*. As in Experiment 1, responses were also scored according to the level of response (1°, 2°, 3°) and according to whether oblique objects and/or particles were also uttered.

Design. For the comprehension task, we employed a one-way design with the between-subjects factor of Age Group (3;5-4;6 vs. 4;9-6;6 vs. 6;10-8;9 vs. adult). The dependent variable was the proportion of trials in which particular manners and/or endstates were chosen, according to the type of contrast being tested: the proportion of trials in which particular manners were chosen in tests of manner sensitivity; the proportion of trials in which particular endstates were chosen in tests of endstate sensitivity; and the proportion of trials in which particular pairings of manner and endstate were chosen in tests of bias. For the production task, we employed a 2 x 3 factorial design with the within-subject factor of Query Topic (content vs. container) and the between-subjects factor of Age Group (3;5-4;6 vs. 4;9-6;6 vs. 6;10-8;9 vs. adult). The dependent variable was the proportion of trials in which either the content or container was encoded as direct object. As in Experiment 1, the performance by the adult group was regarded as the standard for the purpose of establishing standard values for the dependent variables of both the comprehension and production tasks.

Results and Discussion

In replication of the results of Experiment 1, we again found that children were willing to produce locatives of *fill* with the content encoded as direct object, and that children (in this experiment, between 3;5 and 6;6) were biased in their interpretation of filling as having more to do with a pouring manner than a full endstate. Unlike Experiment 1, however, we were also

able to show that children between 4;9-8;9 (and adults) were *sensitive* to the pouring manner of filling. By combining the results of the manner sensitivity and bias tests, we claim that children between 3;5 and 6;6 are likely to have incorporated the pouring manner into the meaning of *fill* given that they are sensitive to the pouring manner, and we found that sensitivity to the pouring manner--within this age range--is associated with instances of syntactic overgeneration. We will discuss, in turn, production results, comprehension results, and tests of association.

Production. In Table 11 we present the proportion of *fill*-content and *fill*-container forms produced as a function of query topic and age group. As in Experiment 1, adults were generally unwilling to produce locatives of *fill* with the content encoded as direct object, with the exception of one adult who uttered one *fill*-content form. By contrast, 34 children (out of 48) produced at least one *fill*-content form (13, 11, and 10 children from the young, mid, and old child groups, respectively). An analysis of variance on the mean proportion of content-locatives produced by subjects of different age groups, and in response to queries of different topic, revealed significant main effects for both age group and query topic. Concerning the main effect of age group, the mean proportions of *fill*-content forms produced by children of increasing age group are 0.50, 0.50, and 0.25, $F(3, 60) = 8.57$, $p < .001$. Planned comparisons (at one-tailed $p < .05$) show not only that the combined group of children ($M = 0.42$) produced significantly more *fill*-content forms than did the adults ($M = 0.02$; $t(62) = 4.22$, $p < .001$), but also that each separate group of children produced significantly more *fill*-content forms than did the adults (for groups of increasing age: $t(30) = 5.08$, $p < .001$; $t(30) = 4.48$, $p < .001$; $t(30) = 3.34$, $p < .002$).

We also found, as in Experiment 1, that the oldest children (here between 6;10-8;9) produced fewer *fill*-content forms than did younger children, this time significantly so ($t(46) = 2.26$, two-tailed $p < .05$). Although differences in the production tasks between experiments

Table 11

Proportion of Trials in which Content- and Container-Locatives of *Fill* were Produced as a Function of Query Topic and Age Group

	AGE GROUP			
	3;5-4;6	4;9-6;6	6;10-8;9	Adult
LOCATIVE FORM				
Content Locatives				
Content-Topic Query	0.53	0.53	0.38	0.03
Container-Topic Query	0.47	0.47	0.12	0.00
Mean	0.50 (22/10/0)	0.50 (27/5/0)	0.25 (16/0/0)	0.02
Container Locatives				
Content-Topic Query	0.38	0.38	0.59	0.97
Container-Topic Query	0.17	0.50	0.88	1.00
Mean	0.42 (. 5/2/0)	0.44 (28/0/0)	0.73 (46/1/0)	0.98

Note: The numerals in parentheses correspond to the frequencies of locatives produced at the 1°/2°/3° level of response. Adults always responded to the primary (1°) query.

prevent a direct comparison of age groups, the fact that older children in both experiments produced fewer *fill*-content forms than younger children probably follows from two sources. First, the older children have *heard* more *fill*-container forms than younger children, which gives the production of *fill*-container forms a "selective advantage" in contexts where either locative form would be appropriate. Second, if the linking of verb syntax and semantics is true, as older children revise their interpretation of *fill*, realizing that no particular manner is incorporated into its meaning (e.g., via observing instances of filling which don't involve a pouring manner), they should also revise their lexical syntactic representation of *fill*, and stop producing *fill*-content forms.

We conclude that children, especially younger ones, overgenerate the content-locative form of *fill*, a fact which bears on whatever lexical representation of syntactic knowledge that children possess of *fill*. That these results cannot be accounted for by any purely non-syntactic explanation--for example, that children responded to the (secondary and tertiary) queries by naming the most salient content or container in the action of filling--can be seen by analyzing separately the responses to the primary query. We find the same results as above: the adults (M = 0.02) produced significantly fewer *fill*-content forms in response to the primary query than did the combined group of children (M = 0.34; $t(62) = 3.80$, $p < .001$), or any of the individual child groups (M_{young} = 0.34, $t(30) = 4.27$, $p < .001$; M_{mid} = 0.42, $t(30) = 3.87$, $p < .001$; M_{old} = 0.25, $t(30) = 3.34$, $p < .002$).

Regarding the main effect of query topic, we found that more *fill*-content forms were produced in response to the content-topic query (M = 0.37) than in response to the container-topic query (M = 0.27), $F(1, 60) = 6.48$, $p < .02$. This is the predicted result given the property of content locatives to treat the content as the topic of discourse (vs. the property of container locatives to treat the container as the topic of discourse). More generally, our new

production task, in which the discourse between experimenter and subject is more structured, appears to be a significant improvement over the task in Experiment 1. We can quantify this improvement by comparing, across experiments, the proportion of all locatives (with either content or container as direct object) which were produced in response to the primary query by children of comparable age--in particular, the 32 younger children in this experiment (3;5-6;6, mean 4;9) and the 32 older children in Experiment 1 (3;6-5;11, mean 4;6). What we find is that significantly more responses were made to the primary query by children in this experiment (M = 0.80) than by children in Experiment 1 (M = 0.52), $t(62) = 3.10$, $p < .005$. Furthermore, we can also assess the improvement in eliciting *full* locative utterances by similarly comparing the proportions of (primary and secondary) locative responses with oblique objects; again, we find significantly more full locatives were produced by children in this experiment (M = 0.63) than by those in the last (M = 0.23), $t(62) = 4.28$, $p < .001$.

A final production result is that children produced content-locatives containing the particle *up* (e.g., *she's filling up grape juice into the glass*), as in Experiment 1, despite the fact that adults reserve that particle for container locatives, apparently to indicate the completeness of a change of state (*John loaded up the wagon with hay*; ?*John loaded up the hay into the wagon*). Of the 48 instances in which children uttered *fill* locative forms with particles, 47 of them involved the use of *up*; of these 47 forms, 15 (.32) were content locatives. Although the willingness of children to utter content locatives with *up* may be routine (simply because they hear *up* in locatives spoken by adults), another explanation is that children interpret the particle *up* in content locative forms much as they would the preposition *up*, according to which the surface of the content rises during the course of the action (e.g., in *she's filling up grape juice into the glass*, *up* refers to the fact that the level of the grape juice rises rather than to the fact that the glass completely changes state).

In summary, our production results provide a straightforward replication of the main finding of Experiment 1--namely, that children make syntactic errors with *fill*, encoding the content rather than the container as its direct object. Our production task also boasts an improved methodology for eliciting full locative forms from young children.

Comprehension. The results of sensitivity and bias tests in the comprehension task provide a detailed record of what subjects take the verb *fill* to mean. We will begin sorting through this record, for subjects of various age, by looking separately at the results of the endstate sensitivity, manner sensitivity, and bias tests. Then we shall put the tests together in an attempt to uncover how the meaning of *fill* changes throughout the course of language acquisition. Although our thesis is that the attribution of a particular manner to the meaning of *fill* licenses the content locative form, we will find a range of meanings attributed to the verb--with different children of the same age (and, by inference, the same child at different times) holding quite different views about its meaning.

Throughout our discussion of the comprehension results, we will make repeated reference to the number of subjects, within an age group, that meet a particular criterion of performance in the comprehension trials. To simplify the presentation of these tallies, and to facilitate their comparison, we have organized all of these counts into one table--Table 12. For each criterion, the underlined numerals indicate that the obtained frequency of subjects is significantly greater than chance, at $p < .05$, according to a two-tailed binomial test. In discussing the particular criteria for sensitivity below, we shall also use *t*-tests to address a different question--whether or not the subjects of an age group, as a whole, are more or less sensitive to an interpretation than the subjects of another age group. In these tests, the dependent variable is the mean difference between the proportions of standard and non-standard responses. The means for

Table 12

Frequency of Subjects Meeting a Criterion of Performance in the Comprehension Task

	AGE GROUP			
CRITERIA	3;5-4;6	4;9-6;6	6;10-8;9	Adult
Endstate Sensitivity				
full > empty (.25,8)	13	10	13	16
empty > full (.25,8)	0*	3	1	0*
full > 3/4-full (.25,8)	8	9	11	15
3/4-full > full (.25,8)	0*	3	2	1
3/4-full > empty (.25,8)	13	11	14	16
empty > 3/4-full (.25,8)	1	2	0*	0*
full > empty, full > 3/4-full, & 3/4-full > empty (.0156,2)	6	9	10	15
full > empty & 3/4-full > empty (.0625,4)	10	10	12	16
Manner Sensitivity				
pouring > dripping (.25,8)	7	9	15	13
dripping > pouring (.25,8)	5	1	0*	0*
Combined Sensitivity (full > 3/4-full) & (pouring > dripping) (.0625,4)	5	5	10	12
~(full > 3/4-full) & (pouring > dripping) (.1875,7)	2	4	5	1
(full > 3/4-full) & ~(pouring > dripping) (.1875,7)	3	4	1	3
~(full > 3/4-full) & ~(pouring > dripping) (.5625,14)	6	3*	0*	0*

Note: The '>' symbol indicates that the left-hand manner or endstate was consistently chosen over the right-hand manner or endstate; the '&' symbol indicates conjunction; the '~' symbol indicates negation. Underlined (asterisked) numerals indicate that the obtained frequency of subjects is significantly greater (lower) than chance, at $p < .05$, according to a two-tailed binomial test. The probability of a single subject meeting a criterion, and the .05 cut-off for a frequency greater than chance, are listed in parentheses following each criterion.

Table 12 (continued)

Frequency of Subjects Meeting a Criterion of Performance in the Comprehension Task

	AGE GROUP			
CRITERIA	3;5-4;6	4;9-6;6	6;10-8;9	Adult
Bias				
full (.0625,4)	7	6	7	16
empty (.0625,4)	0	1	0	0
pouring (.0625,4)	5	6	3	0
dripping (.0625,4)	0	0	0	0
full & pouring (.125,6)	4	0	6	0
full & dripping (.125,6)	0	1	0	0
empty & pouring (.125,6)	0	0	0	0
empty & dripping (.125,6)	0	0	0	0
unbiased (.25,8)	0	2	0	0
Mutually Inconsistent Results				
full bias & ~(full > empty) (.0469,4)	1	0	0	0
empty bias & ~(empty > full) (.0469,4)	0	0	0	0
pouring bias & ~(pouring > dripping) (.0469,4)	1	1	0	0
dripping bias & ~(dripping > pouring) (.0469,4)	0	0	0	0
Mutually Consistent Results				
full bias & (full > empty) (.0156,2)	6	6	7	16
empty bias & (empty > full) (.0156,2)	0	1	0	0
pouring bias & (pouring > dripping) (.0156,2)	4	5	3	0
dripping bias & (dripping > pouring) (.0156,2)	0	0	0	0

Note: The '>' symbol indicates that the left-hand manner or endstate was consistently chosen over the right-hand manner or endstate; the '&' symbol indicates conjunction; the '~' symbol indicates negation. Underlined (asterisked) numerals indicate that the obtained frequency of subjects is significantly greater (lower) than chance, at $p < .05$, according to a two-tailed binomial test. The probability of a single subject meeting a criterion, and the .05 cut-off for a frequency greater than chance, are listed in parentheses following each criterion. See text for an explanation of the bias criteria.

Table 13

Difference between the Proportions of Standard and Non-Standard Interpretations as a Function of Sensitivity Test and Age Group

	AGE GROUP				
	3;5-4;6	4;9-6;6	6;10-8;9	Combined Children	Adult
SENSITIVITY TEST					
Endstate					
full - empty	0.81	0.44	0.75	0.67	1.00
full - 3/4-full	0.50	0.38	0.56	0.48	0.88
3/4-full - empty	0.75	0.56	0.88	0.73	1.00
full - 3/4-full, given full > empty & 3/4-full > empty	0.60	0.90	0.83	0.78	0.88
Manner					
pouring - dripping	0.12	0.50	0.94	0.52	0.81

Note: The '-' symbol indicates that the right-hand (non-standard) manner or endstate was subtracted from the left-hand (standard) manner or endstate in calculating the difference between the proportions of standard and non-standard interpretations. The '>' symbol indicates that the left-hand manner or endstate was consistently chosen over the right-hand manner or endstate; the '&' symbol indicates conjunction.

each relevant criterion, as a function of age group, are listed in Table 13. Each criterion will be discussed in turn.

Endstate Sensitivity. Let's begin with the results of the endstate sensitivity tests, and in particular, with a comparison of adult and child performance on each of the contrasts. Recall that subjects were forced to choose the instance of filling from between three distinct pairs of endstates, the three possible endstates being an empty container, a full container, and a 3/4-full container. In Table 12 we present the tallies of subjects within each age group who, on both trials of a given contrast, chose one particular endstate over another. As Table 12 shows, a significant number of children, in each age group, preferred the standard (adult) endstate to the non-standard endstate, for each of the three contrasts (taken separately); that is, a significant number of children chose the full container over the empty container, the full container over the 3/4-full container, and the 3/4-full container over the empty container. Although the interpretation of these results depends upon the particular contrast under consideration, the binomial tests indicate that the standard endstate interpretation of filling is *accessible* to the children of every age group.

We took the contrast between a full and empty container to be a "liberal" test of a subject's sensitivity to the endstate of filling: the contrasted actions differ in the entire *accomplishment* of filling--not only in the actual *achievement* of filling (i.e., the endstate of the container per se), but also in the *process* of filling (see Vendler, 1967). Thus, a subject may prefer the full container over the empty one because the content and container bear a certain spatial relation to one another throughout the course of the action, regardless of what the final state of the container happens to be (e.g., the level of content rises in the container; the "content-*up*" interpretation). As Table 12 shows, a significant number of subjects in every age group consistently preferred the full containers over the empty ones, with only four children (and

none of the adults) having the opposite preference. This finding is hardly surprising, given the extreme differences between the endstates. More surprising is that the mid-aged children (4;9-6;6) appear to be relatively less sensitive to endstate (on this liberal construal) than are the younger or older subjects. A *t*-test on the mean difference between the proportions of 'full' and 'empty' responses reveals that the mid-aged children ($M = 0.44$) were less sensitive to endstate than were the adults ($M = 1.00$), $t(30) = 2.76$, two-tailed $p < .01$. We also found that the children of the combined group ($M = 0.67$) were less sensitive to endstate than were the adults ($t(62) = 2.10$, $p < .05$), and that the youngest children ($M = 0.81$) were marginally less sensitive to endstate than were the adults ($t(30) = 1.86$, $p < .08$). No other differences between groups were found. See Table 13 for the relevant means.

We considered the contrast between the full and 3/4-full container to be a "conservative" test of endstate sensitivity; the contrasted actions differ in whether or not the endstate of filling is achieved, but not in the amount of content explicitly transferred to the container. Notice that we cannot rule out the possibility that a subject may prefer the full container over the 3/4-full container because the level of content is ultimately higher in the full container, regardless of the state of the container per se. (Children sensitive to *that* interpretation would choose (e.g.) a 1/2-full *tall* glass over a full *short* glass, provided that the level of content in the former is higher than in the latter.) In our estimation, however, the actual endstate of filling is more salient in this contrast than in the liberal test above. For this reason, we expected, and indeed found, that fewer subjects in every age group were sensitive to endstate on this construal (see Table 12). That subjects had greater difficulty with this contrast than with the first (and the third, below) suggests that the endstate of filling is more difficult for children to grasp than is the process of filling. Crucially, we claim that the difficulty in choosing between a full and a 3/4-full container is not simply *perceptual*; our use of color and our protocol (of talking subjects through the drawings) insures that if subjects were inclined to look for a difference, they'd find one.

Across age groups (see Table 13), we found that the mean difference between the proportions of 'full' and '3/4-full' responses is greater for the adults (M = 0.88) than for the youngest children (M = 0.50; $t(30) = 2.09, p < .05$), the mid-aged children (M = 0.38; $t(30) = 2.11, p < .05$), or the combined group of children (M = 0.48, $t(62) = 2.13, p < .05$).

The third contrast, between the 3/4-full container and the empty container, was intended to be a test of whether or not a subject interpreted filling as a process; the contrasted actions differ in the process of filling, but not in the achievement of filling. The results for this test were virtually the same as for the contrast between the full and empty container. In particular, adults and children were willing to choose the 3/4-full container over the empty container despite the incompleteness of the event of filling. (On the other hand, the preference for the 3/4-full container did not seem to carry any implication, at least for the adults, that filling has no definite endpoint.) We also found that the mid-aged children were less sensitive to the process of filling than were the adults, as gauged by the mean difference between the proportions of '3/4-full' and 'empty' responses across groups (M_{adult} = 1.00, M_{mid} = 0.56, $t(30) = 2.41, p < .05$; see Table 13).

Finally, we can look at the performance of subjects in complex tests--combinations of the three endstate contrasts. As expected, the number of subjects who consistently chose the full container over the empty container, the full container over the 3/4-full container, *and* the 3/4-full container over the empty container was significantly greater than chance for each age group (the chance probability of such an outcome for a single subject is .016, and the .05 cut-off is two subjects (out of 16); see Table 12 for the obtained frequencies). In other words, a significant number of subjects, in every age group, appear to view filling as a true accomplishment. Notice that the number of subjects meeting these joint criteria increases with age, contrary to the dip in performance exhibited by the mid-aged children in the separate

full/empty and 3/4-full/empty contrasts. Furthermore, the mid-aged children perform *no worse* than the youngest children even if we put aside the contrast between the full and 3/4-full container, and use the joint criteria of consistently choosing the full container over the empty container and the 3/4-full container over the empty container (see Table 12). We have no explanation for why the mid-aged children performed worse than the younger children on the individual contrasts, but we take the joint test to be a better measure of any true developmental differences.

To address the question of whether the achievement of filling is learned after the process of filling, we performed the following conditional test: for those subjects who consistently chose the full and 3/4-full containers over the empty containers (i.e., given that subjects of an age group were sensitive to the process of filling), is the mean difference between the proportion of 'full' and '3/4-full' responses significantly greater than zero? What we found was that process-sensitive subjects of every group were significantly sensitive to the achievement of filling as well (M_{yng} = 0.60, $t(9) = 3.67$, $p < .01$; M_{mid} = 0.90, $t(9) = 9.00$, $p < .001$; M_{old} = 0.83, $t(11) = 7.42$, $p < .001$; M_{adt} = 0.88, $t(15) = 7.00$, $p < .001$). Across groups of process-sensitive subjects, however, we found that the youngest children were less sensitive to the endstate of filling (in the conservative sense) than the other groups, though this difference failed to reach statistical significance (M = 0.87 (for older children and adults), $t(46) = -1.73$, $p < .09$). By way of contrast, the converse conditional test--given that subjects were sensitive to the achievement of filling, were they sensitive to the process of filling--can be answered positively without qualification; only two children out of 28 (cf. 11/38 for the above test) were sensitive to the achievement, but not process, of filling.

We can summarize the results of our endstate sensitivity tests as follows: The interpretation of filling as an accomplishment--complete with process and achievement--is accessible to the

children of all age groups. Furthermore, it is undoubtedly the case that sensitivity to the accomplishment of filling increases with age, though our own evidence in this regard is inconclusive. With respect to the achievement of filling, however, we found that every group of subjects had greater difficulty in judging the contrast between full and 3/4-full containers than either of the other contrasts, and that young children who were sensitive to the process of filling were somewhat less able, than older children and adults, to choose a full container over a 3/4-full container in picking out the endstate of filling. These latter findings suggest that the process of filling is understood more easily, and before, the endstate of filling.

Manner Sensitivity. In testing for manner sensitivity, our main concern is simply whether subjects preferred one manner of filling to another; that is, did subjects consistently choose the pouring manner over the dripping manner, or the dripping manner over the pouring manner? A second concern is *when* children become sensitive to a manner interpretation of filling.

In Table 12 we've tabulated the number of subjects, per age group, who were consistent in choosing one manner or the other of filling. What we find is that a significant number of the mid-aged children, oldest children, and adults were sensitive to the pouring manner of filling, but not to the dripping manner of filling. The youngest children had access neither to the pouring manner nor to the dripping manner. (In addition, the number of youngest children who were sensitive to any manner--pouring or dripping--failed to reach significance.) Across groups, comparing the mean difference between the proportions of pouring and dripping responses, we found significant differences between the youngest children ($M = 0.12$) and the adults ($M = 0.81$, $t(30) = 2.83$, $p < .01$), between the youngest children and the oldest children ($M = 0.94$, $t(30) = 3.53$, $p < .002$), and between the mid-aged children ($M = 0.50$) and the oldest children ($t(30) = 2.57$, $p < .02$). These means are provided in Table 13.

In interpreting these results, we must remind the reader of an obvious limitation to the sensitivity test--that it doesn't distinguish properties of an action that are essential to the meaning of a verb from those that are merely typical. Thus, the fact that 13 adults consistently chose the pouring manner over the dripping manner in this test only implies that pouring is regarded as a better means, than dripping, to the end of filling, not that pouring is essential to the action of filling. Indeed, only three adults (out of 16) completed both manner sensitivity trials without any reservation; the remainder either hesitated or made some commentary, implying that the choice between pouring and dripping was less than essential. Nonetheless, our results show that the choice was not arbitrary, but presumably reflects a fact about the world to which adults and older children are sensitive. An even stronger claim--that children attribute a particular manner to the meaning of *fill* itself--awaits the results of the bias tests below.

The results also provide us with some clues as to *when* children become sensitive to the manner of filling. One conclusion is that sensitivity to the manner of filling does not necessarily accompany the early use of the verb (i.e., by three-year olds), but usually comes later; in fact, the oldest children are significantly more sensitive to the pouring manner than are either the youngest or mid-aged children. It seems likely, therefore, that the child must know *something* about the verb before a typical manner is potentially incorporated into the meaning of the verb.

This conclusion is in accord with Bowerman's (1982) observation that young children are sensitive to the *type* of arguments that a verb takes (for a similar argument, see Wanner and Gleitman, 1982). As noted in the introduction, Bowerman argues that errors like *I ate my spoon against the pudding* and *I read Mary with a book* do not occur because *eat* and *read* do not take contents and containers as arguments--they are not, in a substantial sense, *locative* verbs. All locative verbs share a concern with potential contents and containers: some are

explicit in encoding the container corresponding to the goal of the transfer (*fill*, *load*); others are explicit in encoding the container corresponding to the source of the transfer (*empty*, *clean*); and still others are non-directional, explicitly encoding the goal or the source or both (*pour*, *dump*). What all of these verbs share, besides their expression in one or both locative forms, is a domain of arguments--an argument space--which specifies (among other things) the shape, size, and dimensionality of potential containers and the mass/count properties of potential contents. Presumably, before a child fixes the meaning of a verb in terms of manner or endstate, he or she must have learned that contents and containers are the appropriate arguments over which to define the meaning of the verb.

On the other hand, it seems plausible to suggest that what the child *doesn't* have to know about the meaning of the verb, before becoming sensitive to manner, is anything about the endstate of the verb (or vice versa). As far as general constraints on verb learning are concerned, the manner with which a content changes location and the endstate of a container are probably independent "dimensions" of verb meaning; there is certainly no principled basis for their dependence. In the case of particular verbs (e.g., *stuff*), however, manner and endstate may be quite interdependent. To test the association of manner and endstate sensitivity with regard to *fill*, we constructed 2 x 2 contingency tables in which subjects were scored as either sensitive or insensitive to the pouring manner and either sensitive or insensitive to the full endstate (depending on their whether or not they always chose the full container over the 3/4-full container). Table 12 lists the number of subjects, within each age group, falling into each of the four cells of the tables. The results of our tests are mixed: we found no evidence of association between manner and endstate sensitivity, for any of the age groups, according to Fisher Exact tests; however, we found that a significant number of subjects in each age group, according to a binomial test, were jointly sensitive to both the pouring manner and the full endstate, but that none of the other joint criteria were satisfied by a significant number of

children. At a minimum, therefore, children within the tested age range are often sensitive to both manner and endstate. Yet we do not make the stronger claim that children *must* become sensitive to the pouring manner and the full endstate at the same time, as if the endstate of filling could only be brought about by the pouring manner.

Bias. Unlike sensitivity tests, bias tests may provide us with a sufficient condition for the essential components of a verb's meaning: if a subject is biased towards an interpretation of a verb's meaning (e.g., identifying the action of filling with a pouring manner, regardless of the co-occurring endstate or the contrasting manner and endstate), then the interpretation is likely to define an essential component of the verb's meaning. Of course, in order to *guarantee* the sufficiency of bias tests, in this respect, we would have to expose subjects to contrasts which vary every dimension that might be relevant to the meaning of a verb; otherwise, we might miss the one essential aspect of its meaning, and our results would reflect merely the most *typical* property, among those tested, of the contexts in which a verb can apply. In this experiment, we have presented subjects with two contrasts, involving variation in two dimensions--the manner in which a substance changes location (pouring or dripping) and the endstate of a container (full or empty).

We interpreted the combined bias tests to yield nine possible results. If a subject consistently chose 'pouring -&- filling' over 'dripping -&- spilling' and 'pouring -&- spilling' over 'dripping -&- filling', he or she was considered biased towards the pouring manner of filling. In a similar way, subjects could display bias towards the dripping manner, the full endstate, or the empty endstate. If a subject split his or her responses on the two presentations of 'pouring -&- spilling' versus 'dripping -&- filling', but consistently chose 'pouring -&- filling' over 'dripping -&- spilling', he or she was considered biased towards the pouring manner *and* the filling endstate (indicating that more tests were needed, or perhaps that the

subject was in the process of switching bias from one component of meaning to the other). Similarly, subjects could display bias towards the pouring manner and the empty endstate, towards the dripping manner and the full endstate, or towards the dripping manner and the empty endstate. Finally, if a subject split his or her responses on the two presentations of each contrast, he or she was considered unbiased.

In Table 12, we've tabulated the number of subjects, per age group, falling into each of the nine outcomes of the bias tests. (For bias towards one component, the chance probability of such an outcome for a single subject is .062, and we'd expect one biased subject (out of 16) by chance; for bias towards two components, the chance probability of such an outcome for a single subject is .125, and we'd expect two biased subjects (out of 16) by chance.) Our results provide a clear replication of the first experiment: a significant number of the youngest (5 out of 16) and mid-aged children (6 out of 16) were biased towards the pouring manner of filling, despite the fact that bias towards the full endstate was also significant for the subjects of each age group (for groups of increasing age: 7, 6, 7, and 16 (adults) identified the action of filling with the full endstate). Interestingly, a significant number of the oldest children (6 out of 16) were biased towards the pouring manner *and* the full endstate, suggesting that they were in a period of transition--perhaps switching their loyalties from the pouring manner to the full endstate. Besides bias towards a pouring manner, a filling endstate, or both, none of the other possible outcomes was achieved by a significant number of subjects.

This selectivity in bias, along with the selectivity in sensitivity reported above, supports our conclusions from the bias tests in the first experiment. In particular, we can safely rule out the possibility that children chose the pouring interpretation of *fill* (regardless of whether they knew its meaning) simply because they liked the panels depicting *pouring*. There are several reasons why this interpretation of the results is untenable: first, we controlled against the

salience of particular panels by using four different picture sets across children. Second--and most importantly--if children performed the forced choice task on the basis of the salience of the pictures, we wouldn't necessarily expect to find any consistent preferences in the drawings that they chose; the salience story doesn't explain why there should be any preference for a *specific* manner at all. Given these results, we must ask: why is the pouring manner so blessed? Why didn't children like the dripping manner? The most plausible answer to these questions is that manner, in general, plays a heavy role in children's representation of verb meaning; and that specific manners--those which make the most "causal" sense--play a heavy role in children's representations of the meanings of specific verbs.

One of the striking differences between the results of the bias and manner sensitivity tests is that bias towards the pouring manner *decreases* significantly for the oldest children whereas sensitivity to the pouring manner *increases* significantly for the oldest children. This combination of results leads us to two conclusions. First, the bias and insensitivity of the youngest children suggests that if young children are sensitive to the pouring manner, they will also be biased towards it. In fact, the ratio of sensitive & biased to sensitive children (for the pouring manner) drops from 0.57 (4/7) to 0.55 (5/9) to 0.20 (3/15) for children of increasing age. Thus, the first assumption that many young children appear to make is that any manner to which they are sensitive must be attributed to the meaning of the verb, an attribution that is presumably a consequence of a general bias towards interpreting verb meaning in terms of the change of location that a substance undergoes. A second conclusion is that the boundary between the oldest and mid-aged children represents a conservative upper ceiling on the willingness of children to incorporate the pouring manner into the meaning of *fill*. This is a *conservative* estimate because the bias tests are jointly sufficient, but not necessary, in determining essential components of meaning; we cannot help but underestimate the number of older children for which the pouring manner is an essential component of the meaning of *fill*.

Thus, a child may regard *both* the pouring manner and full endstate as essential, but when forced to choose, sides with the full endstate or splits his response between the two. Simply put, we have no test which tells us when--in the course of language acquisition--a component of meaning stops being essential to the meaning of a verb, and becomes merely typical of the contexts to which a verb applies. Given this proviso, we can combine our bias and manner sensitivity results to estimate the age range during which children misinterpret the meaning of *fill* as having something essential to do with the (pouring) manner: we estimate that children between 3;5 and 6;6 are likely to have incorporated the pouring manner into the meaning of *fill*, given that they are sensitive to the pouring manner.

Finally, as a check on the reliability of our comprehension data, we performed a series of binomial tests on the number of subjects, per age group, whose combined comprehension results were mutually inconsistent. All of these tests involve outcomes in which subjects exhibited bias towards an interpretation, but no sensitivity towards that interpretation. (We rejected tests in which subjects were sensitive to both manner or endstate, in which case any bias might reasonably occur; we also employed the liberal criterion of endstate (in)sensitivity, since the 3/4-full endstate, being omitted from the bias test, does not bear on the issue of inconsistency.) The results of these tests are given in Table 12, where we find that only three subjects (out of 64) were unreliable. None of the tests reached significance, according to the binomial distribution, and eliminating these three subjects has no appreciable effect on any of our findings. By way of contrast, we also list, in Table 12, the number of subjects whose combined responses *were* mutually consistent. Notice that we find significant numbers of children, per age group, for every outcome involving bias towards, and sensitivity to, the full endstate or the pouring manner.

Summary of Comprehension Results. We can summarize our main findings concerning comprehension as follows. First, children of every age group were sensitive to the endstate of filling, though the process of filling appears to be understood more easily, and before, the endstate of filling. Second, children were also delayed in becoming sensitive to the manner of filling; the youngest children were insensitive to the manner of filling, whereas the older children (4;9-8;9) and adults were sensitive to the *pouring* manner of filling. Third, bias towards the pouring manner was exhibited by a significant number of younger children (3;5-6;6), but not by a significant number of the oldest children, who instead were significantly biased towards the pouring manner *and* the full endstate. In addition, a significant number of subjects of every age group were biased towards the full endstate of filling. We interpreted the results of the bias tests to support the view that different children of the same age may assign different weights to the manner and endstate components of verb meaning, and that some of the oldest children may be in the process of changing their views on the meaning of *fill*. Fourth, the combined sensitivity and bias tests suggest that the pouring manner is an essential component of the meaning of *fill* for those children between 3;5 and 6;6 who are sensitive to the pouring manner.

Given these results, we can offer a rough sketch of how the meaning of *fill* changes throughout the course of development, and ultimately stabilizes as an adult intuition. We assume that early on (usually before the age of our youngest children), children learn that filling is an action which involves contents and containers--that is, which cares about the shape and dimensionality of containers and the mass/count properties of contents. Once a child learns that *fill* is a locative verb, in this sense, he or she may explore several independent interpretations of *fill*. One of the first *specific* facts about filling that children learn is that the action is a (directional) *process*; a specified content always moves into a specified container. Although this information must ultimately be incorporated into the endstate interpretation--

because the accomplishment of filling entails the directional process of filling--there may be a delay before the child understands that the endstate of filling is a necessary part of the meaning of the verb. In fact, it reasonable to suppose that some children may temporarily adopt a "content-up" interpretation--that filling has something essential to do with raising the level of content in a container.

A second fact about filling that some children learn is that the action involves a pouring manner. Of course, other children may never incorporate the pouring manner, or any other manner, into the meaning of the verb. But it is clear from our results that a great number of children, roughly four to seven years of age, may take both the pouring manner and the full endstate to be essential parts of the meaning of *fill*. Thereafter, children gradually become biased away from the pouring manner, and eventually, aware of the fact that the pouring manner isn't a necessary part of the action. This shift is accomplished, we believe, partly on the basis of more exposure to contexts of the verb's usage (some of which don't involve pouring), and partly on the basis of the positive input of the container locative form of *fill* and semantically related verbs (e.g., *cover*; see the General Discussion).

Association. In this section, we will test the following prediction: that the attribution of a particular manner to the meaning of *fill* licenses the content locative form. As in Experiment 1, we're only interested in cases of novel (non-standard) usage, since the association of *standard* syntax and semantics may be transmitted to the child on the basis of positive input. The association of syntactic and semantic errors, on the other hand, implies the use of linking regularities. Unlike Experiment 1, we predict that sensitivity to, and not bias towards, the pouring manner will be the best predictor of *fill*-content production.

A 2 x 2 contingency table was constructed, with each child (per age group) contributing one score to the table. On the semantic dimension, a child was scored as sensitive to the

pouring manner if he or she chose pouring over dripping in both of the manner sensitivity tests; otherwise, a child was scored as insensitive to the pouring manner. On the syntactic dimension, a child was scored as either producing at least one *fill*-content form, or producing no *fill*-content forms. To our surprise, no significant associations were found for the separate child groups or for the combined child group (3;5-8;9), although the Fisher Exact Test for the mid-aged children approached significance ($p < .08$). The contingency tables for each child group and for the combined child groups are presented in Table 14.

An examination of Table 14 gives us some insight into these results. We find that sensitivity to the pouring manner of *fill* was a reliable predictor of *fill*-content production for the youngest and mid-aged children, but not for the oldest children; of the 16 children between 3;5 and 6;6 who were sensitive to the pouring manner, 15 produced at least one *fill*-content form, whereas only 9 out of the 15 oldest children (who were sensitive to pouring) produced at least one *fill*-content form. Why is it the case that six of the oldest children produced no content locatives, despite the fact that they were sensitive to the pouring manner? For whatever reason, we put aside the data for the oldest children and performed a post hoc test on the association of syntax and semantics for the combined group of youngest and mid-aged children. This time, we found a significant association according to a Fisher Exact Test (one-tailed $p < .02$): a child between 3;5 and 6;6 who is sensitive to the pouring manner of filling will tend to produce the *fill*-content form, whereas a child between 3;5 and 6;6 who is not sensitive to the pouring manner of filling will not tend to produce the *fill*-content form.

In retrospect, this result should not have been surprising. Returning to the question of why six of the oldest children produced no content locatives, despite the fact that they were sensitive to the pouring manner, we can offer two independent explanations based on our comprehension and production results. First, for three of these subjects (and the one mid-aged

Table 14
Contingency Tables per Age Group on the Association between Sensitivity to the Pouring Manner in the Interpretation of *Fill* and the Occurrence of *Fill*-Content Locatives

3;5-4;6:

	Produced At Least One Content Locative	Produced No Content Locatives
Pouring Sensitive	7	0
Pouring Insensitive	6	3

4;9-6;6:

	Produced At Least One Content Locative	Produced No Content Locatives
Pouring Sensitive	8	1
Pouring Insensitive	3	4

6;10-8;9:

	Produced At Least One Content Locative	Produced No Content Locatives
Pouring Sensitive	9	6
Pouring Insensitive	1	0

YOUNGEST AND MID-AGED CHILDREN (3;5-6;6):

	Produced At Least One Content Locative	Produced No Content Locatives
Pouring Sensitive	15	1
Pouring Insensitive	9	7

COMBINED CHILDREN (3;5-8;9):

	Produced At Least One Content Locative	Produced No Content Locatives
Pouring Sensitive	24	7
Pouring Insensitive	10	7

Note: Each child counted in a table contributed one score to that table. A subject was scored as sensitive to the pouring manner if he or she always chose pouring over dripping in the sensitivity tests.

child who performed similarly), it is arguable that they regarded the pouring manner as typical, rather than essential, of filling. This claim is consistent with the fact that none of these subjects were biased towards the pouring manner, and with our earlier estimate (based on the combined bias and manner sensitivity tests) that the pouring manner is an essential component of the meaning of *fill* for those youngest and mid-aged children who are sensitive to it. We should thus expect the strongest result of association for children in this age range. A second explanation, discussed in Experiment 1, is simply that the older children have heard more utterances of the *fill*-container form, which would selectively discourage older children from uttering *fill*-content forms despite any influence of linking In fact, the oldest children produced only half as many content locatives as either the mid-aged or the youngest children.

Although we interpret the results of our tests to provide strong support for the hypothesis that children make use of linking regularities in order to produce non-standard forms, it does not seem to be the case that a child who is insensitive to the pouring manner of filling *must* produce only the *fill*-container form. Overall, of the 17 children between 3;5 and 8;9 who were insensitive to the pouring manner, 10 produced at least one *fill*-content form. We must explain why these ten children appear to have violated the predicted linking of syntax and semantics. Our answer now, as in Experiment 1, is that we have tested merely *one* misinterpretation of filling; it is very plausible that some of these children are sensitive to *another* interpretation of filling which licenses the content-locative form--for example, the content-*up* interpretation. (Of the ten children who produced at least one *fill*-content form, but were insensitive to the pouring manner, four produced at least one content locative with the particle *up*.)

Finally, we can point to one clear improvement in our present association test over the association test in Experiment 1: the criterion of sensitivity picks out more of the children who

Table 15
Contingency Tables per Age Group on the Association between Bias in the Interpretation of *Fill* and the Occurrence of *Fill*-Content Locatives

3;5-4;6:

	Produced At Least One Content Locative	Produced No Content Locatives
Pouring Bias	5	0
Full Bias	6	1

4;9-6;6:

	Produced At Least One Content Locative	Produced No Content Locatives
Pouring Bias	5	1
Full Bias	4	2

6;10-8;9:

	Produced At Least One Content Locative	Produced No Content Locatives
Pouring Bias	1	2
Full Bias	4	3

COMBINED CHILDREN (3;5-8;9):

	Produced At Least One Content Locative	Produced No Content Locatives
Pouring Bias	11	3
Full Bias	14	6

Note: Each child counted in a table contributed one score to that table. A subject was scored as biased towards an interpretation if he or she always chose drawings consistent with that interpretation in the bias tests.

produced *fill*-content forms than did the criterion of bias (without also picking out more of the younger children who produced only *fill*-container forms). In particular, of the 34 children who produced at least one *fill*-content form, only 10 (a miss ratio of .29) were insensitive to the pouring manner; in the bias association test of Experiment 1, by contrast, 11 out of 19 children (who produced *fill*-content forms) were biased towards the full endstate (miss ratio = .58). Furthermore, we found much the same result in this experiment: when contingency tables were constructed with semantic levels corresponding to bias towards the pouring manner versus bias towards the full endstate (see Table 15), 14 out of 25 children (who produced *fill*-content forms) were biased towards the full endstate (miss ratio = .56). As expected on the basis of the selective shift of frequency in the cells of these tables, we found no association of syntax and semantics according to Fisher Exact tests.

In conclusion, the evidence from our "case study" of *fill* strengthens our conclusions from Experiment 1: that children make syntactic errors in verb learning; that children make semantic errors in verb learning; and that the syntactic and semantic errors are associated with one another. In particular, we have shown that children between 3;5 and 6;6 who are sensitive to the pouring manner of filling are likely to have incorporated that manner into the meaning of the verb, and we found that sensitivity to the pouring manner--within this age range--is associated with production of the *fill*-content form. We conclude that misinterpretations of particular verbs, coupled with universal linking regularities, may account for the occurrence of the syntactic errors. More generally, by improving upon the method in Experiment 1, we were able to understand more about how the meaning and syntax of a verb changes throughout the course of development.

Experiment 3

Experiments 1 and 2 focused on English locative verbs in the process of being acquired. A complimentary approach, which we used in the following four experiments, is to look at the "fast mapping" of novel (i.e., made-up) verbs: we taught children and adults novel verbs in a neutral syntax (e.g., *this is mooping*), and then tested their willingness to encode the content or container as the direct object of the verb. By varying whether, and how, the content or the container is affected in the meaning of a verb, we assessed whether, and under what circumstances, subjects can induce the syntactic privileges of a verb on the basis of its meaning. We predict that children and adults should produce relatively more content locatives for those verbs in which the content changes location in a particular manner, and relatively more container locatives for those verbs in which the container changes state in a particular way.

Method

Subjects. Forty-eight children and sixteen adults, all native speakers of English living in the Boston area, participated in the study. The children fell equally into three age groups, roughly comparable to those of Experiment 2: sixteen between 3;4 and 4;5 (mean 3;11); sixteen between 4;7 and 5;11 (mean 5;1); and sixteen between 6;5 and 8;6 (mean 7;5). (Eight children, who failed to understand the taught verbs or were otherwise confused, distracted, or shy, were replaced in the design.) The children were drawn from middle-class day-care and after-school programs in Newton, Needham, and Watertown. The adults were MIT students, ranging in age from roughly 18 to 22 years, and were paid for their participation.

Materials. In the priming/pretesting phase of the experiment, we used a cup and some (loose) marbles. In the remainder of the experiment, we endeavored to discourage subjects

from making rote responses by using two separate pairs of materials, each consisting of a surface/container and some content: an 8" square piece of felt and a sealed packet (clear plastic bag) of pennies; and an 8" square piece of plastic and a sealed packet of marbles. During the teaching and testing phases of the experiment, the current surface was placed on a wooden stand, which was constructed so that one side provided support for the entire surface whereas the other side provided support only for the perimeter of the surface.

Two verb meanings were created using the same pairs of materials: in the *manner* condition, a packet of content was moved to a (supported) surface in a zig-zagging manner; in the *endstate* condition, a surface came to sag down, as the result of the content being moved (in a nondescript fashion) to the *un*supported surface. By using the same pairs of materials for both actions (within subject), we insured that any difference in a subject's performance for these actions was not due to the salience of the materials themselves.

Corresponding to these two novel actions were two stems, *pilk* and *keat*. The combination of meanings and stems into verbs was counterbalanced across subjects in an age group.

Procedure. The procedure consisted of a priming/pretesting phase, a teaching phase for each novel verb, and a production test for each novel verb. Subjects were tested in a single (20-minute) session by two experimenters (one eliciting responses; the other observing) in an area as free as possible from potential distractions. For the children, each novel verb was introduced as a puppet word by a puppet, the second word being introduced by a new puppet.

After being introduced to all of the physical materials in the study (except the stand), subjects were pretested on, and primed with, locative forms of *pour* and *fill*. The pretesting consisted of asking subjects (unfocused) queries about actions of pouring and filling, and recording whether they encoded the content or container as the direct object of the verb. The

syntactic priming consisted of subsequently giving subjects *feedback* on the pretesting--modeling the full content locative of *pour* and the full container locative of *fill*. We included pretesting as a miniature replication of the production test in Experiments 1 and 2; we included the priming to help orient our subjects towards the domain of locative verbs.

For example, the experimenter would use the following script: "do you know the word *fill*? ... when I do this (as experimenter moves marbles, a few at a time, into a cup) ... and it ends up like that (as experimenter finishes filling the cup with marbles) ... it's called *filling*." The experimenter presented this action and utterance three times. After the third presentation, the experimenter asked the subject, "using the word *fill*, can you tell me what I'm doing?" As in Experiments 1 and 2, we tested production at three levels of response, if necessary: in those cases where the query failed to elicit an unambiguous direct object, we followed up with the secondary prompt "filling what?"; in those cases where the secondary prompt failed to elicit an unambiguous response, we followed up with the tertiary prompt "filling the cup or filling the marbles?" (with the order of choices balanced across subjects in an age group). Regardless of the subject's final response, the experimenter modeled a full container locative of *fill*, "I'm filling the cup with marbles," and had the subject repeat the sentence. The pretesting and priming were performed for both *pour* and *fill*, with order balanced across subjects within an age group.

Each subject was then taught two novel verbs. one specifying a particular manner (zig-zagging) and the other specifying a particular endstate (sagging). The verbs were taught and elicited one at a time, with the order of verb meaning balanced across subjects in an age group. The experimenter began by introducing (or having a puppet introduce) the phonetic form of the verb: "can you say *keat* (*pilk*)?... say *keat* (*pilk*)." Thereafter, the teaching phase consisted of showing the subject, and having the subject act out, positive and negative instances of the

verb's meaning. In a typical round of teaching, for the endstate verb, the experimenter would say to a subject: "let me show you what keating is ... when I do this (as experimenter moves a packet of marbles towards an unsupported piece of plastic in a nondescript manner) ... and it ends up like that (as experimenter places the packet of marbles onto the piece of plastic, causing the plastic to sag) ... it's called *keating*." The experimenter then repeated this positive illustration once, before giving a *negative* model: "now let me show you something that's *not* keating ... when I do this (as experimenter moves a packet of marbles towards a *supported* piece of plastic) ... and it ends up like that (as experimenter places the packet of marbles onto the piece of plastic, the end result being *no* change in the shape or conformation of the surface) ... it's *not* called keating." After thus presenting the subject with two positive and one negative example of keating, the experimenter then asked the subject to act out one positive ("can you show me what keating is?) and one negative ("can you show me something that's not keating?") instance of keating. If children failed to grasp the meaning of the verb, the experimenter repeated a positive model and a positive comprehension query, using the same pair of materials. We note that the use of negative models and queries was rarely a source of confusion, and generally helped children to focus on the intended regularity in manner or endstate.

For the manner meaning, the same procedure for teaching was used. In the positive model of the manner meaning, the experimenter moved a packet of content towards the surface in a zig-zagging manner, ultimately placing the content onto the (supported) surface; in the negative example, by contrast, the content was moved in a *bouncing* manner. One difference between the manner and endstate conditions involved the linguistic context of the two verb types. For the endstate verb, the linguistic context read, "when I do this ... and it ends up like that ..."; for the manner verb, "when I do this ... and it ends up *over there* ..." In both cases, we used the linguistic context to encourage children to consider the whole of the action (cf. the linguistic

context for describing the two-panel drawings in Experiment 1). The difference between the two conditions, and in particular the use of "it ends up *over there*", was prompted by the propensity for children (in pilot testing) to disregard the involvement of the surface/container when acting out the manner verb. (In retrospect, it may have been wiser to use the same linguistic context in both conditions, but we note that both the content and container are referred to in the manner script--hence, we did not introduce a *non-semantic* bias for the subject to associate the manner action with either the content or container. Of course, inducing in the subject a *semantic* construal of the event--for example, that the marbles move to the surface in a particular manner--was precisely the purpose of the teaching phase.)

The teaching protocol, as described above, was repeated for the second pair of materials (i.e., the packet of pennies and the piece of felt). Both pairs of materials were used in the teaching and testing of each verb meaning, with the sequence of materials switched for the second verb (within subject) and balanced across subjects within an age group. (Note: several of the children offered spontaneous definitions for the verbs, which inspired us to elicit "definitions" after the second round of comprehension queries (i.e., by asking, "why is this (not) keating?") from 20 of the children. The descriptions of the endstate verb were often informative, and will be discussed below.)

After teaching a novel verb to a subject, we tested the ability of the subject to produce locative forms with that verb. The elicitation technique here was similar to that of Experiment 1. (Experiments 3 and 4 predate the revised technique used in Experiment 2.) Subjects were asked two focused queries, one which focused the content (in one pair of materials) and one which focused the container (in the other pair of materials). To make sure that a subject was attending to the novel action with a particular pair of materials, the experimenter always preceded a production query with a final positive comprehension query. Thus, after finishing

the teaching phase with the second pair of materials, the experimenter would switch back to the original pair of materials, run a final comprehension check ("can you show me what keating is?"), and then pose (e.g.) a content-topic query: "what are these called? ... (experimenter waits for, or supplies if necessary, the response of *marbles*); say *keating* ... (experimenter waits for response); can you tell me, with the word *keating*, what I'm doing with the *marbles* (as experimenter performs action)?" The experimenter would then re-introduce the second set of materials with a quick comprehension test, and pose the other (e.g., container-topic) query. As in the pretesting and previous experiments, we tested production at three levels of response, if necessary. (The order of choices in the tertiary query was balanced within subject.)

The second verb was taught and tested with the same protocol as the first. Besides switching the order of material pairs for the second verb, we also switched the order of query topics so that the full sequence of query topics was either content-container-container-content or container-content-content-container. Both of these switches, in unison, guaranteed that the *same two items* (i.e., marbles and felt or pennies and plastic) were focused for both verbs, within subject. This is an important precaution: because of the lack of any syntactic information about the new verb available to a subject, we wanted to make sure that the focusing of different materials (with potentially different salience) in the production task could not account for any differences in a subject's performance for the two verb meanings. Furthermore, the combination of verb meaning, query-topic order, and material order was counterbalanced *across* subjects within an age group.

Scoring. The responses to the pretesting and testing were scored according to whether the direct object corresponded to the content or container in the performed action. We used the conventions and protocol in the Scoring Section of Experiment 1 for determining whether or not the use of a pronoun (e.g., *filling it*) was ambiguous. Responses which were

undecipherable or which included no specification of the direct object were coded as *other*. As in Experiment 1, responses were also scored according to the level of response (1°, 2°, 3°) and according to whether oblique objects and/or particles were also uttered.

Unlike the previous experiments with English verbs, children in this experiment were much more prone to resort to familiar terms when confronted with a novel action and verb. We therefore recorded any relevant speech uttered during the session, including any (spontaneous or elicited) descriptions of the novel actions offered by the subjects during the teaching phase of the experiment, and any spontaneous substitution of English verbs for the novel ones being elicited during the testing phase of the experiment.

Design. We employed a 2 x 2 x 4 factorial design with the within-subject factors of Verb Meaning (manner vs. endstate) and Query Topic (content vs. container), and the between-subjects factor of Age Group (3;4-4;5 vs. 4;7-5;11 vs. 6;5-8;6 vs. adult). The dependent variable was the proportion of trials in which either the content or container was encoded as direct object.

Results and Discussion

In discussing our results, we will address two questions. First, did children and adults produce relatively more content locatives for the novel manner verb, and relatively more container locatives for the novel endstate verb? Second, did children generally prefer the content locative to the container locative (as they did in Experiment 1), and if so, why?

The proportion of trials in which subjects produced content and container locatives as a function of verb meaning, query topic, and age group are presented in Table 16. A measure of greater transparency (and less redundancy), however, is the *difference* between these proportions for a given subject. Unlike the elicitation of genuine English non-alternators, we

are not interested here in the occurrence or non-occurrence of incorrect forms. Why not? Don't we predict that the manner verb should only be produced in the content-locative form, and that the endstate verb should only be produced in the container-locative form? Ideally, yes--but it must be recognized that the demands of this experiment are far from ideal. We are asking subjects to use a verb which they have never heard modeled before. Furthermore, we are expecting a syntactic difference between two verbs taught back-to-back to the same subject with the same materials. For these reasons, and because our production task approximated a forced choice procedure, we put more faith in a measure of relative production that we have called (in Experiment 1) the *preference score*--the proportion of trials in which a content locative is produced minus the proportion of trials in which a container locative is produced. Preference scores range from +1.0 (a strong preference for content locatives) to –1.0 (a strong preference for container locatives). Mean preference scores are listed in Table 17 as a function of verb meaning, query topic, and age group.

We performed an Analysis of Variance on the mean preference score, with the within-subject variables of Verb Meaning and Query Topic and the between-subjects variable of Age Group. We found significant main effects for verb meaning and query topic, a marginal interaction between verb meaning and query topic, and a marginal main effect for age group. We'll discuss these findings, and relevant follow-up analyses, in turn.

On the basis of the main effect of verb meaning, we can answer our first question in the affirmative: subjects had a significantly greater preference score (i.e., produced relatively more content locatives than container locatives) for the novel manner verb (M = 0.77) than for the novel endstate verb (M = 0.34), F(1, 60) = 21.50, $p < .001$. This main effect was also observed for responses to the primary query (M_{mnr} = 0.30, M_{end} = 0.12, $F(1, 60) = 6.50$, $p < .02$), arguing that the effect is a reflection of genuine syntactic knowledge. Within age group,

Table 16

Proportion of Trials in which Content and Container Locatives of Novel Verbs were Produced as a Function of Meaning, Query Topic, and Age Group

MEANING-FORM	AGE GROUP			
	3;4-4;5	4;7-5;11	6;5-8;6	Adult
Manner				
Content Locatives				
Content-Topic Query	1.00	1.00	1.00	0.88
Container-Topic Query	0.88	0.94	0.69	0.62
Mean	0.94 (6/24/0)	0.97 (8/23/0)	0.84 (17/10/0)	0.75 (20/4/0)
Container Locatives				
Content-Topic Query	0.00	0.00	0.00	0.06
Container-Topic Query	0.12	0.06	0.31	0.31
Mean	0.06 (2/0/0)	0.03 (0/1/0)	0.16 (5/0/0)	0.19 (6/0/0)
Endstate				
Content Locatives				
Content-Topic Query	0.94	0.88	0.75	0.69
Container-Topic Query	0.56	0.69	0.38	0.44
Mean	0.75 (7/17/0)	0.78 (12/13/0)	0.56 (11/7/0)	0.56 (16/2/0)
Container Locatives				
Content-Topic Query	0.06	0.12	0.19	0.31
Container-Topic Query	0.38	0.31	0.62	0.56
Mean	0.22 (4/3/0)	0.22 (3/4/0)	0.41 (10/3/0)	0.44 (14/0/0)

Note: The numerals in parentheses correspond to the frequencies of locatives produced at the 1°/2°/3° level of response.

Table 17

Mean Preference Score for Novel Verbs as a Function of Meaning, Query Topic, and Age Group

	AGE GROUP			
MEANING	3;4-4;5	4;7-5;11	6;5-8;6	Adult
Manner				
Content-Topic Query	1.00	1.00	1.00	0.81
Container-Topic Query	0.75	0.88	0.38	0.31
Mean	0.88	0.94	0.69	0.56
Endstate				
Content-Topic Query	0.88	0.75	0.56	0.38
Container-Topic Query	0.19	0.38	-0.25	-0.12
Mean	0.53	0.56	0.16	0.12

Note: Mean preference score was calculated by subtracting the mean proportion of trials in which container locatives were produced from the mean proportion of trials in which content locatives were produced.

we found a significant effect of verb meaning for the mid-aged children ($M_{mnr} = 0.94$, $M_{end} = 0.56$, $F(1, 15) = 5.87$, $p < .03$), the oldest children ($M_{mnr} = 0.69$, $M_{end} = 0.16$, $F(1, 15) = 6.31$, $p < .03$), and the adults ($M_{mnr} = 0.56$, $M_{end} = 0.12$, $F(1, 15) = 5.79$, $p < .03$), and a marginally significant effect of verb meaning for the youngest children ($M_{mnr} = 0.88$, $M_{end} = 0.53$, $F(1, 15) = 3.85$, $p < .07$). (The smaller effect of verb meaning for the youngest children is the artifact of a ceiling effect; see the analysis by verb order below.) The effect also held for the responses to the primary query for the oldest children (marginally: $M_{mnr} = 0.38$, $M_{end} = 0.03$, $F(1, 15) = 4.13$, $p < .06$) and adults ($M_{mnr} = 0.44$, $M_{end} = 0.06$, $F(1, 15) = 4.66$, $p < .05$).

These results show that subjects of every age used the meanings of the novel verbs to predict a difference in their syntactic privileges. In fact, we can cite a strong reason why the present findings undoubtedly *underestimate* the ability of English speakers to make use of linking regularities. Although the design of this experiment is powerful in that verb meaning is varied within subject, the drawback is that strong *set effects* influenced the behavior of the subjects. In the light of the similarity of the actions and materials, it is remarkable that our subjects overcame the great temptation either to respond consistently with the content (or container) or to repeat back the topic of the query in a mechanical fashion. The systematic deviation from these patterns that we found is difficult to explain without invoking the linking regularities. But furthermore, we can document the influence of the set effects by analyzing the production of novel verbs as a function of verb order; set effects should wash out the effects of verb meaning for the *second* verb taught and tested. (Note: in the analysis of the initial two trials, Verb Meaning becomes a between-subjects factor.) And this is precisely what we find: for the manner verb, the preference score shows a marginally significant *drop* from the first order ($M = 0.88$) to the second order ($M = 0.66$, $F(1, 56) = 3.85$, $p < .06$); whereas for the

endstate verb, the preference score shows a significant *rise* from the first order (M = 0.16) to the second order (M = 0.53, $F(1, 56) = 4.26$, $p < .05$). Thus, the effect of verb meaning is reduced from a highly significant value for the first order ($M_{difference}$ = 0.72, $F(1, 56)$ = 21.59, $p < .001$) to a nonsignificant value for the second order ($M_{difference}$ = 0.12, $F(1, 56)$ = 0.73, $p \approx .40$). In addition, we note that the preference score for the manner verb in the first order was *at ceiling* (M = 1.00) for the youngest and mid-aged children, accounting for the apparent reduction in the effect of verb meaning for these children. Mean preference scores are listed in Table 18 as a function of verb meaning, query topic, age group, and verb order.

The fact that there is a highly significant effect of verb meaning (on initial as well as combined trials) shows that children and adults can use affectedness to predict *some* syntactic difference between verbs. We would like to demonstrate, moreover, that the observed syntactic difference corresponds to the discrete syntactic judgments that adults ultimately make about non-alternating verbs like *pour* and *fill*. In other words, although we don't predict--for reasons given above--that the manner and endstate verbs should *only* be produced in content and container locatives, respectively, our case would be strengthened if the manner and endstate verbs were produced *more often* in content and container locatives, respectively. The problem here, however, is that subjects did *not* prefer to link the container argument of the endstate verb to the direct object position. An examination of Tables 17 and 18 reveals that the mean preference score for the endstate verb was negative only once--in initial trials for the adult group (M = -0.38); and even here, this preference score was not significantly less than zero according to a two-tailed t-test against zero ($t(7) = -1.43$, $p \approx .20$). Otherwise, the preference score was consistently positive in initial (and combined) trials. In a series of planned two-tailed t-tests, no other preference scores for the endstate verb in initial trials were significantly

Table 18

Mean Preference Score for Novel Verbs as a Function of Meaning, Query Topic, Age Group, and Verb Order

	AGE GROUP			
	3;4-4;5	4;7-5;11	6;5-8;6	Adult
1ST ORDER				
Manner Meaning				
Content-Topic Query	1.00	1.00	1.00	0.75
Container-Topic Query	1.00	1.00	0.75	0.50
Mean	1.00	1.00	0.88	0.62
Endstate Meaning				
Content-Topic Query	0.75	0.50	0.62	0.00
Container-Topic Query	0.38	0.25	-0.50	-0.75
Mean	0.56	0.38	0.06	-0.38
2ND ORDER				
Manner Meaning				
Content-Topic Query	1.00	1.00	1.00	0.88
Container-Topic Query	0.50	0.75	0.00	0.12
Mean	0.75	0.88	0.50	0.50
Endstate Meaning				
Content-Topic Query	1.00	1.00	0.50	0.75
Container-Topic Query	0.00	0.50	0.00	0.50
Mean	0.50	0.75	0.25	0.62

Note: Mean preference score was calculated by subtracting the mean proportion of trials in which container locatives were produced from the mean proportion of trials in which content locatives were produced.

different from zero. (Note that the preference score was significantly greater than zero in *combined* trials for the youngest children ($M = 0.53$, $t(15) = 3.44$, $p < .005$) and mid-aged children ($M = 0.56$, $t(15) = 3.09$, $p < .01$); however, the analysis of responses in the initial trials is obviously more appropriate in tests of absolute preference against zero, given the set effects discussed above). By contrast, all of the preference scores for the *manner* verb in initial (and combined) trials were significantly greater than zero: $M_{yng,1st} = 1.00$; $M_{mid,1st} = 1.00$; $M_{old,1st} = 0.88$, $t(7) = 7.00$, $p < .001$; $M_{adt,1st} = 0.62$, $t(7) = 2.38$, $p < .05$).

Why is it the case that subjects did not prefer the container-locative form of the endstate verb? Our endstate condition--in which the experimenter placed the content onto an unsupported surface, causing the surface to sag--was chosen because it seemingly involves the particular way in which a surface changes state, but not the particular manner in which some content changes location. By hypothesis, then, the container-locative form of the endstate verb should be preferred. We have independent reasons to believe, however, that subjects were sometimes confused by the complexity of the endstate action, and that some of them may have attributed a particular *manner* to the "endstate" verb. (Of course, we would be guilty of circular reasoning if we simply took the preference score as a metric of which entity is affected in the meaning of a verb.) One likely source for a manner interpretation of the sagging action quickly became apparent to us during the course of our testing: although we designed the content to sink into the unsupported surface by virtue its weight alone, the experimenter often had to *nudge the content* into the unsupported surface in order to initiate the sagging. Our guess is that children were as sensitive to this impetus as they were to anything else about this complex action. (In fact, the interaction of the content and the container in this action probably makes it more akin to *stuffing* than to *sagging*).

In support of the claim that the endstate action was too complex and equivocal in interpretation, we analyzed children's responses in those 16 cases (out of 48) where they gave some unambiguous indication (via the utterance of English verbs, nouns, or adjectives) of whether they were interpreting the sagging action in terms of the content, the container, or both. (We were fastidious about eliminating from consideration references to actions such as *putting*, *going down*, and even *squishing*, which are vague or ambiguous in what they could tell us about the child's interpretation of the novel verb.) Of these 16, we found that 10 focused on the content--most often, whether the content (with or without the container) was going or hanging down ($M_{10} = 0.55$). Of the remaining 6 children, most interpreted the action in terms of the shape of the container (e.g., the plastic was "bent down" or "made crinkly"; $M_6 = 0.17$). In addition, several of the children were distracted by the stand used in performing this action. By contrast, we found much less variation in the children's interpretation of the manner action. Of the 21 children who provided commentary on the manner action, they invariably characterized it as involving some manner in which the *content* moved--such as waving, zig-zagging, swinging, or shaking (versus hopping, jumping, or bouncing for the negative model).

Because the stuffing/sagging action was subject to various interpretations--involving the affectedness of either the content or the container--it is therefore quite plausible, albeit post-hoc, to suppose that subjects should have shown no absolute preference for either the content- or container-locative form of the verb. In addition, we suggest that the effect of verb meaning might have been even greater had we invented a purer, simpler endstate verb. In Experiment 4, we shall take up the challenge of inventing such a verb, and thereby attempting to show an absolute preference for the container-locative form of the verb, as well as a larger effect of verb meaning.

Returning to our analysis of variance, we also found a highly significant main effect of query topic, according to which subjects produced relatively more content locatives when the content was the topic of the query (M = 0.80) than when the container was the topic of the query (M = 0.31, $F(1, 60)$ = 32.50, $p < .001$). This effect shows, as in our previous experiments, the predictable influence of discourse topic on the subjects' choice of direct object. We also found an interaction between query topic and verb meaning, indicating that the effect of query topic is greater in the sagging/stuffing condition than in the manner condition. Although this interaction was only marginally significant in the original ANOVA ($p < .09$), it was highly significant in an ANOVA on initial (first order) trials (for the manner condition, $M_{content\text{-}topic}$ = 0.94, $M_{container\text{-}topic}$ = 0.81; for the sagging/stuffing condition, $M_{content\text{-}topic}$ = 0.47, $M_{container\text{-}topic}$ = -0.16; $F(1, 56)$, $p < .01$). As these preference scores show, this interaction is the consequence of a ceiling effect.

The last finding in our original ANOVA is a main effect of age group, marginally significant at $p < .07$ ($F(3, 60) = 2.57$), which indicates that subjects of different age produced different proportions of content and container locatives: for subjects of increasing age, the mean preference scores were 0.70, 0.75, 0.42, and 0.34 (adults), respectively. Treating the adult preference score as the standard, we performed a series of two-tailed t-tests on the difference in preference score between age groups. We found that the youngest ($t(30) = 1.88$, $p < .07$) and mid-aged children ($t(30) = 2.00$, $p < .06$), but not oldest children, produced (marginally) relatively more content locatives than the adults. Virtually identical t-test results (rounded to hundredths) were found for the youngest and mid-aged children when tested against the oldest children, indicating that the oldest children patterned with the adults. The finding that the youngest and mid-aged children prefer the content locative, relative to oldest children and adults, was confirmed for the initial trials, where the range of preference scores is broader, and the trend for younger subjects to have greater preference scores appears to be

roughly linear (with respect to the mean age of the children): for groups with increasing mean ages of 3;11, 5;1, 7;5, and adult, the mean preference scores were 0.78, 0.69, 0.47, and 0.12, respectively ($F(3, 56) = 3.55, p < .02$). Again treating the adult preference score (based on initial trials) as the standard, we found that the youngest ($t(28) = 2.90, p < .01$) and mid-aged children ($t(28) = 2.28, p < .05$), but not oldest children, produced significantly more content locatives than the adults. We also performed a contrast analysis on the hypothesis that the mean preference score of the child groups is an inverse linear function of mean age. The weights used in the analysis were calculated from the (unrounded) mean ages to be 3 (youngest), 1 (mid-aged), and -4 (oldest). The test results were not significant ($F(1, 42) = 1.26, p > .25$).

These results bear directly on our second question, of whether children generally preferred the content locative to the container locative: children between 3;4 and 5;11 did indeed show a preference for content locatives relative to older children and adults. This finding is of course reminiscent of the overgeneration of the *fill*-content form that we observed in Experiments 1 and 2, and raises the question of how the same children performed in the pretesting. The results of the pretesting, presented in Table 19, replicate the relevant finding from Experiment 1: that children are prone to overgenerate the content locative of *fill*, but not the container locative to *pour*. Eighteen children out of 48 (38%) produced the *fill*-content form (11, 3, and 4 from groups of children of increasing age), whereas *none* of the children produced the *pour*-container form. Adults made no errors with either *pour* or *fill*, as expected. The proportion of trials in which children made syntactic errors with *fill* is comparable in this pretesting (M = 0.38) to the error rates from Experiments 1 ($M = 0.47$) and 2 ($M = 0.42$); the higher yield in those extended tests is undoubtedly due to the focusing provided by the queries (versus none in the pretest).

Table 19

Proportion of Pretest Trials in which Content and Container Locatives of *Pour* and *Fill* were Produced as a Function of Age Group

VERB-FORM	AGE GROUP			
	3;4-4;5	4;7-5;11	6;5-8;6	Adult
Pour				
Content Locatives	1.00 (2/14/0)	1.00 (5/11/0)	1.00 (9/7/0)	1.00 (15/1/0)
Container Locatives	0.00	0.00	0.00	0.00
Fill				
Content Locatives	0.69 (4/7/0)	0.19 (1/2/0)	0.25 (1/3/0)	0.00
Container Locatives	0.31 (2/3/0)	0.81 (8/5/0)	0.75 (9/3/0)	1.00 (15/1/0)

Note: The numerals in parentheses correspond to the frequencies of locatives produced at the 1°/2°/3° level of response.

Finally, let us consider the question of *why* the younger children show a preference for content locatives relative to older children and adults. At the outset, we can at least rule out any account based *solely* on the distribution of locatives in the language (e.g., Bowerman, 1982). On such an account we would not expect the effect of verb meaning that we found; children of a certain age would most plausibly overgeneralize the dominant form in a domain (e.g., content locatives) on the basis of a shallow, easily learnable semantic distinction, such as that between contents and containers. They would not necessarily be sensitive to finer-grained semantic distinctions which pertain to the affectedness of an entity in the meaning of a verb. (Strictly speaking, the distributional account is *consistent* with the data presented here if one makes the unlikely assumption that the effect of age group and the effect of verb meaning arise from two conflicting mechanisms for lexicosyntactic productivity--involving the distribution of forms in a language and Object Affectedness, respectively.)

Alternatively, we shall assume that *one* mechanism for productivity--involving the linking rule of Object Affectedness--*must* be used to explain the verb meaning effect. Once this mechanism is assumed, we shall argue (as we have heretofore) that the younger children's relative preference for content locatives falls out of general considerations of the nature of the child's developing cognitive and perceptual systems. Specifically, a perceptual/cognitive *manner bias*, in conjunction with Object Affectedness, might be the source of both the overgeneralization of content locatives to established verbs such as *fill* and the extension of content locatives to novel verbs such as our zig-zagging and sagging/stuffing inventions. The crucial similarity between filling and stuffing/sagging, we think, is that in both cases there is a "nucleus" of a manner for children to latch onto; as we saw in the case of *fill*, this nucleus may be a characteristic, not essential, component of meaning in the adult's semantic representation of the verb. Similarly, younger children may have been especially sensitive to the (regular)

stuffing manner of the sagging/stuffing action, and on the basis of the linking rule, produced content locatives of the corresponding verb. Their relative preference for content locatives, therefore, would result from the production of content locatives for both novel verbs. In favor of this explanation, we note that adults and oldest kids in this experiment pattern together--reminiscent of our finding in experiment 2 that bias begins to fade, at least with respect to *fill*, around the age of 7.

In summary, we found that children and adults, in using verbs which they have never heard modeled before, show systematic differences in the syntactic privileges that they assign to the verbs: they produced relatively more content locatives for the manner verb, designed to specify the particular manner in which a content changed location, and relatively more container locatives for the endstate verb, designed to specify the particular endstate of a container which resulted from an action. We concluded that the ability of subjects to predict these syntactic privileges implicated the linking regularity of Object Affectedness. One problem with this interpretation, however, was that subjects did not show an absolute preference in linking the container argument of the endstate verb to the direct object position. In this case, we argued that the endstate action was too complex and equivocal in interpretation, possibly resulting in the lack of a clear syntactic preference for the verb, and we predicted that the teaching of a purer, simpler endstate verb might enable subjects to choose its container-locative form with more consistency. Finally, we found that younger children preferred content locatives relative to older children and adults. We argued that this result was similar to the selective overgeneration of the content locative to *fill* in the pretesting and in earlier experiments, and we suggested that a general manner bias, in conjunction with Object Affectedness, might be the source of both syntactic preferences.

Experiment 4

The primary purpose of this experiment is to teach children and adults a "pure" endstate verb--one in which subjects could not possibly infer a particular manner from the means by which the endstate of a container is achieved. We present each subject with an action in which the change of state of a container is not only very salient (i.e., a color change), but also "causally isolated" from the preceding change in the location of the content (i.e., the proximal cause of the state change is chemical, and thus not observable). If the means of a salient state change were opaque, we reasoned, subjects would be unable or unwilling to infer a particular manner from the presentation of the action. Consequently, we predict that subjects should show an absolute preference for the container-locative form of the endstate verb, as well as a larger effect of verb meaning.

In addition, we take a new, complimentary approach to the details of design. The major changes include: the use of interchangeable contents and containers (so that we could balance their pairing); the use of two possible manners and endstates (so that the identification of affected entities would not be susceptible to properties peculiar to any one manner or endstate); and the focusing of each material (content or container) once and only once during each session (so that we might lessen the influence of set effects due to focusing the same two materials for both verbs, as in Experiment 3).

Method

Subjects. Forty-eight children and sixteen adults, all native speakers of English living in the Boston area, participated in the study. The children fell into three age groups corresponding closely to those in Experiment 3: sixteen between 3;5 and 4;5 (mean 3;10); sixteen between 4;7 and 5;8 (mean 5;1); and sixteen between 6;7 and 8;5 (mean 7;3). (One child was replaced in the

design for being unresponsive in the production task; three children were replaced in the design due to experimenter error, or to a difficulty with the materials.) The children were drawn from middle-class day-care and after-school programs in Cambridge, Newton, Needham, and Watertown. The adults were MIT undergraduates, ranging in age from roughly 18 to 22 years, and were paid for their participation. (One adult was replaced in the design due to his color-blindness.)

Materials. As in Experiment 3, each subject made use of two separate sets of materials, although in this experiment the pairing of interchangeable objects (contents) and surfaces (containers) was balanced across subjects in an age group. The surface was either a 2-3" x 4-5" piece of (absorbent) paper or a 2-3" x 4-5" piece of felt; the object was either a 1" square piece of sponge or a cotton ball. The materials being currently used were always damp to the touch: the surface was saturated with cabbage juice; the object was saturated with either water, lemon juice, or a baking-soda solution (in water). (Before each performance of an action, the surface was placed in a tray in order to contain the liquids. After each performance, the used object and surface were discarded, the tray was wiped dry, and a new pair of materials was introduced, depending upon the next procedure.) In addition, we used a cup and some marbles, as in Experiment 3, in the priming/pretesting phase of the experiment. (Note: the priming/pretesting materials were not used again in the main conditions, as they were in the previous experiment; we thereby precluded the (slight) possibility of capitalizing on the heightened salience of the marbles in subsequent production testing.)

Two verb meanings were created using pairs of materials. In the endstate condition, the surface changed color in an acid-base reaction from purple (the color of unadulterated cabbage juice) to either pink (when an object saturated with lemon juice was moved to the surface) or green (when an object saturated with baking-soda solution was moved to the surface). Within

subject, the endstate was always the same; across subjects, the endstate was pink as often as it was green. In the *manner* condition, an object was moved to a surface in a particular manner, either zig-zagging or bouncing. Within subject, the manner was always the same; across subjects, the manner was zig-zagging as often as it was bouncing. The surface in this condition was saturated with cabbage juice and the object was saturated with water; no color change resulted from their interaction. As in the previous experiment, our use of the same pairs of materials for both actions (within subject) insured that any difference in a subject's performance for these actions was not due to the salience of the materials themselves. One modification here is our use of two possible manners and endstates across subjects, so that the identification of affected entities would not be susceptible to properties peculiar to any one manner of motion or color change. In addition, we counterbalanced the four possible pairings of objects and surfaces (paper-sponge; paper-cotton ball; felt-sponge; felt-cotton ball) with verb meaning so that each pairing of object and surface occurred as often in the manner condition as it did in the endstate condition, across the subjects in an age group.

Corresponding to these two novel actions were two stems, *moop* and *keat*. We thought that it might be easier for young children to pronounce *moop* and to keep it distinct from *keat* (versus *pilk*, used in Experiment 3). The combination of meanings and stems into verbs was counterbalanced across subjects in an age group.

Procedure. The procedure used in this experiment is virtually the same as that of Experiment 3. It consisted of a priming/pretesting phase, a teaching phase for each novel verb, and a production test for each novel verb. Subjects were tested in a single (20-minute) session by two experimenters (one eliciting responses; the other observing) in an area as free as possible from potential distractions. For the children, each novel verb was introduced as a puppet word by a puppet, the second word being introduced by a new puppet.

After being introduced to all of the physical materials in the study (*not* including the cabbage juice, the water, the lemon juice, the baking-soda solution, or the tray), subjects were pretested on, and primed with, locative forms of *pour* and *fill*. The pretesting consisted of asking subjects (unfocused) queries about actions of pouring and filling, and recording whether they encoded the content or container as the direct object of the verb. The syntactic priming consisted of subsequently giving subjects *feedback* on the pretesting--modeling the full content locative of *pour* and the full container locative of *fill* . For example, the experimenter would use the following script: "do you know the word *fill*? ... when I do this (as experimenter moves marbles, a few at a time, into a cup) ... and it ends up like that (as experimenter finishes filling the cup with marbles) ... it's called *filling*." The experimenter presented this action and utterance three times. After the third presentation, the experimenter asked the subject, "using the word *fill*, can you tell me what I'm doing?" As in the previous experiments, we tested production at three levels of response, if necessary. Regardless of the subject's final response, the experimenter modeled a full container locative of *fill*, "I'm filling the cup with marbles," and had the subject repeat the sentence. The pretesting and priming were performed for both *pour* and *fill*, with order balanced across subjects within an age group.

As in Experiment 3, each subject was then taught two novel verbs, one specifying a particular manner (e.g., zig-zagging) and the other specifying a particular endstate (e.g., pink). (The *un*specified manner and endstate for a subject (cf. bouncing and green) were used in the negative models.) The verbs were taught and elicited one at a time, with the order of verb meaning balanced across subjects in an age group. The experimenter began by introducing (or having a puppet introduce) the phonetic form of the verb: "can you say *keat* (*moop*)? ... say *keat* (*moop*)." Thereafter, the teaching phase consisted of showing the subject, and having the subject act out, positive and negative instances of the verb's meaning. For example, the

experimenter would say to a subject, "let me show you what keating is ... when I do this (as experimenter moves a cotton ball, saturated with lemon juice, towards a piece of paper in a nondescript manner) ... and it ends up like that (as experimenter pats the paper with the cotton ball, causing the paper to change color from purple to pink) ... it's called *keating*." The experimenter then repeated this positive illustration once (with a fresh pair of materials), before giving a *negative* model: "now let me show you something that's *not* keating ... when I do this (as experimenter moves a cotton ball, saturated with *baking-soda solution*, towards a piece of paper in a nondescript manner) ... and it ends up like that (as experimenter pats the paper with the cotton ball, causing the paper to change color from purple to *green*) ... it's *not* called keating." After thus presenting the subject with two positive and one negative example of keating, the experimenter then asked the subject to act out one positive ("can you show me what keating is?) and one negative ("can you show me something that's not keating?") instance of keating. If subjects failed to grasp the meaning of the verb, the experimenter repeated a positive model and a positive comprehension query, using a fresh pair of materials. We note that subjects were quite sensitive to the specified color change, and that they were generally willing to attend to the endstate action without knowing the underlying "kitchen chemistry" (the curious were promised, and all were given, an explanation at the close of the session).

For the manner meaning, the same procedure for teaching was used. In the positive model of the manner meaning, the experimenter moved an object towards the surface in a (e.g.) zig-zagging manner, ultimately placing the object onto the surface; in the negative example, by contrast, the object was moved in a (e.g.) *bouncing* manner. As in Experiment 3, the linguistic context for the manner verb read, "when I do this ... and it ends up *over there* ..."

The teaching protocol, as described above, was repeated for the second pairing of materials (i.e., the piece of sponge and the piece of felt). Both pairs of materials were used in the

teaching and testing of each verb meaning, the sequence of materials for the first verb being counterbalanced with the sequence of materials for the second verb (across the subjects within an age group using a particular pairing of object and surface).

After teaching a novel verb to a subject, we tested the ability of the subject to produce locative forms with that verb. The elicitation technique here was identical to that of Experiment 3. Subjects were asked two focused queries, one which focused the object (in one pair of materials) and one which focused the surface (in the other pair of materials). As in the previous experiment, the experimenter always preceded a production query with a final positive comprehension query. After finishing the teaching phase with the second pair of materials, the experimenter would switch back to the original pair of materials, run a final comprehension check ("can you show me what keating is?"), and then pose (e.g.) a content-topic query: "what is this called? ... (experimenter waits for, or supplies if necessary, the response of *a cotton ball*); say *keating* ... (experimenter waits for response); can you tell me, with the word *keating*, what I'm doing with the *cotton ball* (as experimenter performs action)?" The experimenter would then re-introduce the second set of materials with a quick comprehension test, and pose the other (e.g., container-topic) query. As in the pretesting and previous experiments, we tested production at three levels of response, if necessary. (The order of choices in the tertiary query was balanced within subject.)

The second verb was taught and tested with the same protocol as the first. Here, we departed from the design of the previous experiment in two ways: first, the order of query topics for the first verb was counterbalanced with the order of query topics for the second verb, across subjects in an age group; second, each material (object or surface) was focused once and only once per session, and across subjects in an age group each material (in a given pairing) was focused an equal number of times within meaning condition. We thought that this design

would rule out the possibility that the focusing of different materials (with potentially different salience) could account for any observed effect of verb meaning, while at the same time lessening the homogenizing influence of focusing the same two materials for both verbs (as in Experiment 3).

Scoring. The responses to the pretesting and testing were scored according to whether the direct object corresponded to the content or container in the performed action. We used the conventions and protocol in the Scoring Section of Experiment 1 for determining whether or not the use of a pronoun (e.g., *filling it*) was ambiguous. Responses which were undecipherable or which included no specification of the direct object were coded as *other*. As in the previous experiments, responses were also scored according to the level of response (1°, 2°, 3°) and according to whether oblique objects and/or particles were also uttered. We also recorded any spontaneous speech uttered during the session, including the substitution of English verbs for the novel ones being elicited, and any commentary by the subjects on the novel verbs during the teaching phase of the experiment.

Design. We employed a 2 x 2 x 4 factorial design with the within-subject factors of Verb Meaning (manner vs. endstate) and Query Topic (content vs. container), and the between-subjects factor of Age Group (3;5-4;5 vs. 4;7-5;8 vs. 6;7-8;5 vs. adult). The dependent variable was the proportion of trials in which either the content or container was encoded as direct object.

Results and Discussion

We will address two questions: First, did children and adults produce more content locatives for the novel manner verb, and more container locatives for the novel endstate verb?

Second, did younger children (again) show a preference for producing content locatives, relative to the oldest children and adults?

In Table 20 we present the proportion of trials in which subjects produced content and container locatives as a function of verb meaning, query topic, and age group. As in Experiment 3, we used these means to derive a more useful dependent measure--the *preference score*: the proportion of trials in which a content locative is produced minus the proportion of trials in which a container locative is produced. Preference scores range from +1.0 (a strong preference for content locatives) to –1.0 (a strong preference for container locatives). Mean preference scores are listed in Table 21 as a function of verb meaning, query topic, and age group.

We performed an Analysis of Variance on the mean preference score, with the within-subject variables of Verb Meaning and Query Topic and the between-subjects variable of Age Group. We found significant main effects for verb meaning and query topic, and a marginal interaction between verb meaning and age group. The main effect of verb meaning indicates that subjects had a significantly higher preference score (i.e., produced relatively more content locatives than container locatives) for the manner verb (M = 0.34) than for the endstate verb (M = -0.83), $F(1, 60) = 106.94$, $p < .001$. The main effect of verb meaning was also observed for responses to the primary query (M_{mnr} = 0.04, M_{end} = -0.38, $F(1, 60) = 26.58$, $p < .001$, confirming that the effect is a reflection of genuine syntactic knowledge. Furthermore, this effect is even greater for responses in the initial trials, which were free from the set effects due to the similarity of the actions and materials: the mean preference score for the manner verb in initial trials is 0.62; the mean preference score for the endstate verb in initial trials is -0.97, $F(1, 56) = 200.08$, $p < .001$. (Note that the set effects are not completely comparable in Experiments 3 and 4; although the second responses always involve a regression towards the

Table 20

Proportion of Trials in which Content and Container Locatives of Novel Verbs were Produced as a Function of Meaning, Query Topic, and Age Group

MEANING-FORM	AGE GROUP			
	3;5-4;5	4;7-5;8	6;7-8;5	Adult
Manner				
Content Locatives				
Content-Topic Query	0.62	0.88	0.81	0.69
Container-Topic Query	0.44	0.62	0.69	0.62
Mean	0.53 (1/16/0)	0.75 (1/23/0)	0.75 (9/15/0)	0.66 (15/6/0)
Container Locatives				
Content-Topic Query	0.38	0.12	0.19	0.31
Container-Topic Query	0.56	0.38	0.31	0.38
Mean	0.47 (5/10/0)	0.25 (1/7/0)	0.25 (4/4/0)	0.34 (11/0/0)
Endstate				
Content Locatives				
Content-Topic Query	0.19	0.00	0.25	0.00
Container-Topic Query	0.12	0.00	0.06	0.00
Mean	0.16 (3/2/0)	0.00	0.16 (2/3/0)	0.00
Container Locatives				
Content-Topic Query	0.75	1.00	0.75	1.00
Container-Topic Query	0.81	1.00	0.94	1.00
Mean	0.78 (5/20/0)	1.00 (5/27/0)	0.84 (14/13/0)	1.00 (29/3/0)

Note: The numerals in parentheses correspond to the frequencies of locatives produced at the 1°/2°/3° level of response.

Table 21

Mean Preference Score for Novel Verbs as a Function of Meaning, Query Topic, and Age Group

MEANING	AGE GROUP			
	3;5-4;5	4;7-5;8	6;7-8;5	Adult
Manner				
Content-Topic Query	0.25	0.75	0.62	0.38
Container-Topic Query	-0.12	0.25	0.38	0.25
Mean	0.06	0.50	0.50	0.31
Endstate				
Content-Topic Query	-0.56	-1.00	-0.50	-1.00
Container-Topic Query	-0.69	-1.00	-0.88	-1.00
Mean	-0.62	-1.00	-0.69	-1.00

Note: Mean preference score was calculated by subtracting the mean proportion of trials in which container locatives were produced from the mean proportion of trials in which content locatives were produced.

mean, the relative size of the regression was skewed towards the content locative in Experiment 3 (the mean preference score, collapsing over verb meaning, rose from 0.52 to 0.59) and towards the container locative in Experiment 4 (the mean preference score dropped from -0.17 to -0.31).) Mean preference scores are listed in Table 22 as a function of verb meaning, query topic, age group, and verb order.

The marginal interaction between verb meaning and age group indicates that the effect of verb meaning was significantly different for different age groups, $F(3, 60) = 2.35$, $p < .09$. A precise interpretation of this interaction is difficult to offer, though it appears as if the effect of verb meaning varies cubically as a function of age group (or quadratically, if we just consider the child groups): for the youngest children, $M_{mnr} = 0.06$, $M_{end} = -0.62$; for the mid-aged children, $M_{mnr} = 0.50$, $M_{end} = -1.00$; for the oldest children, $M_{mnr} = 0.50$, $M_{end} = -0.69$; for the adults, $M_{mnr} = 0.31$, $M_{end} = -1.00$. In any case, we do not attach too much importance to this trend, for two reasons. First, the interaction of verb meaning and age group was not found for responses in the initial trials ($p > .25$). Second, despite any variation across age groups, the effect of verb meaning was significant within each age group: for groups of increasing age, $F_{yng}(1, 15) = 6.51$, $p < .025$; $F_{mid}(1, 15) = 90.00$, $p < .001$; $F_{old}(1, 15) = 27.21$, $p < .001$; $F_{adt}(1, 15) = 30.77$, $p < .001$.

The main effect of query topic indicates that subjects had a significantly higher preference score for responses to the content-topic query ($M = -0.13$) than for responses to the container-topic query ($M = -0.35$), $F(1, 60) = 10.00$, $p < .005$. (For responses in the initial trials, $M_{content\text{-}topic} = -0.08$, $M_{container\text{-}topic} = -0.27$, $F(1, 56) = 4.85$, $p < .05$.) This is expected, given the discourse function of content locatives to treat the content as the topic of conversation, and the discourse function of container locatives to treat the container as the topic

Table 22

Mean Preference Score for Novel Verbs as a Function of Meaning, Query Topic, Age Group, and Verb Order

	AGE GROUP			
	3;5-4;5	4;7-5;8	6;7-8;5	Adult
1ST ORDER				
Manner Meaning				
Content-Topic Query	0.75	0.75	1.00	0.75
Container-Topic Query	0.00	0.50	0.75	0.50
Mean	0.38	0.62	0.88	0.62
Endstate Meaning				
Content-Topic Query	-0.88	-1.00	-1.00	-1.00
Container-Topic Query	-0.88	-1.00	-1.00	-1.00
Mean	-0.88	-1.00	-1.00	-1.00
2ND ORDER				
Manner Meaning				
Content-Topic Query	-0.25	0.75	0.25	0.00
Container-Topic Query	-0.25	0.00	0.00	0.00
Mean	-0.25	0.38	0.12	0.00
Endstate Meaning				
Content-Topic Query	-0.25	-1.00	0.00	-1.00
Container-Topic Query	-0.50	-1.00	-0.75	-1.00
Mean	-0.38	-1.00	-0.38	-1.00

Note: Mean preference score was calculated by subtracting the mean proportion of trials in which container locatives were produced from the mean proportion of trials in which content locatives were produced.

of conversation. We also found, for responses in the initial trials, an interaction between verb meaning and query topic, indicating that the effect of query topic was greater for responses to the manner verb ($M_{content\text{-}topic} = 0.81$, $M_{container\text{-}topic} = 0.44$) than for responses to the endstate verb ($M_{content\text{-}topic} = -0.97$, $M_{container\text{-}topic} = -0.97$, $F(1, 56) = 4.85$, $p < .05$). Notice that this interaction is due to a floor effect, equal (but opposite) to the ceiling effect observed in Experiment 3 (esp., in the interaction of verb meaning and query topic observed there).

The main effect of verb meaning replicates our main finding from Experiment 3: that children and adults can use the meanings of verbs to predict syntactic differences between them. Moreover, it appears as though the effect of verb meaning is greater in this experiment than in the previous one. To quantify this increase, we pooled the data from Experiments 3 and 4, and ran an analysis of variance with the within-subject variable of Verb Meaning and the between-subjects variable of Experiment (Experiment 3 vs. Experiment 4). We found a highly significant interaction between experiment and verb meaning, indicating that the effect of verb meaning is greater in this study ($M_{mnr} = 0.34$, $M_{end} = -0.83$, $M_{difference} = 1.17$) than in the last ($M_{mnr} = 0.77$, $M_{end} = 0.34$, $M_{difference} = 0.42$), $F(1, 126) = 26.00$, $p < .001$. This interaction is also highly significant for the initial trials: $M_{diff,exp4} = 1.59$, $M_{diff,exp3} = 0.72$, $F(1, 124) = 19.63$, $p < .001$. As to *why* the effect of verb meaning is greater in this study than in the last, we observe that the increase is entirely due to a shift in the preference scores for the endstate verb. Looking primarily at responses in initial trials (on the grounds that the influence of set effects on the preference scores for second trials is not comparable across experiments), we found that subjects had a significantly lower preference score for the coloring verb ($M_{exp4} = -0.97$) than for the stuffing/sagging verb ($M_{exp3} = 0.16$), $t(62) = 7.61$, $p < .001$; by way of contrast, the preference score for the manner verb was only marginally

different between the two experiments, and in the wrong direction (to account for the increased effect of verb meaning found here), $M_{exp4} = 0.62$, $M_{exp3} = 0.88$, $t(62) = 1.91$, $p < .07$. Furthermore, we found that subjects of every age group had a significantly lower preference score for the coloring verb than for the stuffing/sagging verb: for the youngest children, $M_{exp4} = -0.88$, $M_{exp3} = 0.56$, $t(14) = 5.02$, $p < .001$; for the mid-aged children, $M_{exp4} = -1.00$, $M_{exp3} = 0.38$, $t(14) = 4.25$, $p < .001$; for the oldest children, $M_{exp4} = -1.00$, $M_{exp3} = 0.06$, $t(14) = 4.43$, $p < .001$; and for the adults, $M_{exp4} = -1.00$, $M_{exp3} = -0.38$, $t(14) = 2.38$, $p < .05$. Significant differences for each age group were also found for the responses in the combined trials.

One of our main goals in performing this experiment was to show that children and adults would have an absolute preference for the container-locative form of the endstate verb if that verb were designed so as to prevent subjects from inferring a particular manner from the presentation of the action. The selective shift between experiments in the production of the endstate verb shows that subjects in this experiment, unlike those in the last, did indeed prefer to link the container argument of the coloring verb to the direct object position. As is clear from Tables 21 and 22, the preference scores for the coloring verb were consistently negative (for initial and combined trials) whereas the preference scores for the manner verb were consistently positive. Planned two-tailed *t*-tests verified that the preference score for the endstate verb in initial trials was significantly lower than zero for the youngest children, as well as the other age groups: $M_{yng} = -0.88$, $t(7) = -7.00$, $p < .001$; $M_{mid} = -1.00$; $M_{old} = -1.00$; $M_{adt} = -1.00$. The preference score for the manner verb in initial trials was significantly greater than zero for all but the youngest children: $M_{yng} = 0.38$, $t(7) = 1.43$, $p \approx .20$; $M_{mid} = 0.62$, $t(7) = 3.42$, $p < .02$; $M_{old} = 0.88$, $t(7) = 7.00$, $p < .001$; $M_{adt} = 0.62$, $t(7) = 2.38$, $p < .05$.

We have shown that children and adults had an absolute preference for the container-locative form of the endstate verb, but not really that we prevented subjects from inferring a particular manner from the presentation of the coloring action. In fact, we found that children were unanimous in their interpretation of the verb. Most revealing is that 33 children (out of 48, or 69%) spontaneously uttered a color name (including *pink*, *green*, *blue*, *yellow*, *magenta*, and *turquoise*!) during the course of learning the novel endstate verb. True to the strongest of our predictions, the mean preference score for these children was -1.00. Furthermore, these children didn't utter color names simply because they had seen the corresponding colors; in the majority of these cases (20), children commented upon an actual *change* in the color of the surface (13 instances; e.g., "it's going green" or "it's turning pink"), or even upon the essential nature of a *specific color change* to the meaning of the endstate verb (7 instances; e.g., "because mooping only turns it that pinkish color"). (In five instances, children also made reference to rubbing the surface in order to change its color; e.g., "rubbing it and it's turning blue.") In contrast, only three children made any reference to the liquids involved (e.g., "where's the paint"), and only one child made reference to an object (e.g., "from the sponge"). We can safely conclude, therefore, that we succeeded in creating a pure endstate verb. Thus, the results of this experiment support our hypothesis that children and adults can predict a difference in the syntax of novel verbs on the basis of their meanings, and moreover that this syntactic difference (in absolute preference) corresponds to the discrete syntactic judgments that adults ultimately make about non-alternating verbs like *pour* and *fill*.

Finally, we turn to our second question: did younger children show a preference for producing content locatives, relative to the oldest children and adults, as they did in Experiment 3? The answer here is uniquivocally *no*. Besides finding no main effect, or even a trend towards a main effect, for age group, we found that the mean preference score for the youngest and mid-aged children (M = -0.27) was roughly comparable to that for the oldest children and

adults (M = -0.22). The same non-result holds for the analysis of responses in the initial trials. By contrast, we note that these same children (especially the youngest) overgenerated the content locative of *fill* in the pretest. The results of the pretesting, presented in Table 23, replicate the relevant findings from Experiments 1 and 3: that children are prone to overgenerate the content locative of *fill*, but not the container locative of *pour*. Seventeen children out of 48 (35%) produced the *fill*-content form (10, 4, and 3 from groups of children of increasing age), whereas only *one* of the (mid-aged) children produced the *pour*-container form. Adults made no errors with either *pour* or *fill*, as expected. The proportion of trials in which children made syntactic errors with *fill* is comparable in this pretesting (M = 0.35) to the error rates from Experiments 1 (M = 0.47) and 2 (M = 0.42) and 3 (M = 0.38); the higher yield in Experiments 1 and 2 is undoubtedly due to the focusing provided by the queries (versus none in the pretest).

This pattern of results is not surprising if we assume that the relative preference for content-locatives in the pretesting, as in Experiment 3, was the result of a heightened sensitivity to manner which could only exert its influence on a *regularity* in the manner of an action. As we suggested in the previous experiment, younger children may have been especially sensitive to the (regular) pouring manner of *fill* and the stuffing manner of the stuffing/sagging action, and on the basis of the linking rule, produced content locatives of those verbs. In the present experiment, however, we successfully designed an endstate verb in which *no* (causally-transparent) regularity could be discerned in the manner of the coloring action. It follows, therefore, that younger children wouldn't be any more prone, than older children or adults, to interpret the endstate verb as specifying a particular manner, and thus that they wouldn't be any more prone to produce content locatives of that verb.

Table 23

Proportion of Pretest Trials in which Content and Container Locatives of *Pour* and *Fill* were Produced as a Function of Age Group

VERB-FORM	AGE GROUP			
	3;5-4;5	4;7-5;8	6;7-8;5	Adult
Pour				
Content Locatives	1.00 (6/10/0)	0.94 (4/11/0)	1.00 (11/5/0)	1.00 (16/0/0)
Container Locatives	0.00	0.06 (1/0/0)	0.00	0.00
Fill				
Content Locatives	0.62 (3/7/0)	0.25 (0/4/0)	0.19 (1/2/0)	0.00
Container Locatives	0.38 (3/3/0)	0.75 (5/7/0)	0.81 (12/1/0)	1.00 (16/0/0)

Note: The numerals in parentheses correspond to the frequencies of locatives produced at the 1°/2°/3° level of response.

In summary, we replicated our main finding from Experiment 3: that children and adults can use the meanings of verbs to predict syntactic differences between them. Furthermore, the effect of verb meaning was greater in this experiment than in the last, attributable to a shift in the mean preference score for the endstate verb: subjects in this experiment, unlike those in the last, preferred to link the container argument of the coloring verb to the direct object position. In light of our observation that the children appeared to be unanimous in their interpretation of the endstate verb as specifying the particular color change of a surface, we took these findings to confirm our prediction that children and adults would have an absolute preference for the container-locative form of the endstate verb if that verb were designed so as to prevent subjects from inferring a particular manner from the presentation of the action. Consistent with this interpretation, as well, was the finding that younger children showed no preference for producing content locatives of the novel verbs, relative to the oldest children and adults, although they did selectively overgenerate the *fill*-content form. In this case, we argued that *fill*--but not our "pure" endstate verb--was susceptible to a general sensitivity to manner which can only exert its influence if a regularity is perceivable and perceived in an action.

Experiment 5

In Experiments 3 and 4, we tested the ability of subjects to predict syntactic differences between verbs on the basis of their meanings. The meanings that we invented bore some correspondence to *pour* and *fill*; one was designed to specify the particular manner in which content changed location (cf. *pour*) and the other was designed to specify the particular change of state of a container (cf. *fill*). In the following two studies, we again address the issue of whether children can use meaning to predict syntax--but this time, we test components of meaning which correspond more closely to the manner and endstate of an alternating verb, such as *stuff*. As we discussed in the Introduction and Experiment 1, the manner and endstate components of alternating verbs are often mutually constraining or interpredictable. For example, in stuffing clothes into a hamper, the clothing must be forced into the hamper (perhaps to the extent that the clothing is compressed) *because* the capacity of the hamper is exhausted; conversely, the fact that the clothing must be forced into the hamper seems to imply that the hamper is already *stuffed* (perhaps to the extent that the hamper bulges).

If the manner and endstate of a verb are interpredictable, the question arises as to what dictates when one locative form or the other will be used on a given occasion. In the case of some alternating verbs, such as *load*, *spray*, and *sprinkle*, it appears as if the meaning of the verb specifies the *potential* for the container to be affected in a particular way, but whether or not the container is actually construed by the speaker/hearer as affected depends upon the *extent* of the action. For example, if John loaded two bullets into a gun, leaving most of the chambers of the gun empty, it would be odd to describe this event with the sentence *John loaded the gun with bullets*. On the other hand, if John repeatedly loaded bullets into the gun until the capacity of the gun was exhausted, its potential as a firearm being fully enabled, then the sentence *John loaded the gun with bullets* would be acceptable (and in fact, more informative than the content

locative *John loaded bullets into the gun*). Thus, the container locative, but not content locative, of the verb *load* carries the necessary implication that the container is totally or holistically affected (Anderson, 1971; Schwartz-Norman, 1976).

A complete account of how locative verbs are mastered must explain not only how children overgeneralize, and ultimately unlearn, forms such as the content locative of *fill*, but also how children come to use the alternative locative forms of a verb such as *load* properly. The linking rule of Object Affectedness can, in principle, account for both phenomena (provided that we can understand what it means for the *potential* for affectedness to be part of the meaning of alternating verbs; see the General Discussion). In Experiments 5 and 6, we test the ability of children and adults to predict the syntactic form of verbs which vary in the extent to which the corresponding actions are performed: in the partitive condition, an action involving the addition of an object (content) to a container (e.g., placing a peg into a hole on a board) is performed once; in the holistic condition, the same action is repeated until the capacity of the container is exhausted (i.e., until the container becomes holistically affected). We predict that children and adults should produce relatively more content locatives for the verb in the partitive condition, and relatively more container locatives for the verb in the holistic condition.

In addition, we varied whether or not subjects were provided with explicit *aspectual cues* concerning the extent of the action. *Aspect* refers to the distribution of an action over time; for example, whether an action is an activity or process with no definite endpoint in time (e.g., pouring, loading bullets), or an *accomplishment*--an action of duration with a definite endpoint in time (e.g., filling, loading a gun) (Vendler, 1967; Tenny, 1988). An aspectual cue, in our sense, is a syntactic frame which accompanies the performance of an action. In particular, we hypothesized that the frame "I'm not done V-ing yet...I'm not done V-ing yet...*now* I'm done V-ing...I V-ed" might provide subjects with the information that an action has a definite

endpoint, at which time (marked by *now*) a container becomes holistically affected. We therefore tested the prediction that children and adults would be more willing to use the holistic verb in the container-locative form after hearing aspectual cues than before hearing them. (Note: we do not propose that children *must* hear such overt syntactic cues before they can learn that a verb specifies a (potential) accomplishment; only that such cues provide a sufficient condition for reaching that conclusion. Furthermore, we assume that the more exposure to instances of a holistic action (with or without aspectual cues) that a child has, the more likely he or she will be to conclude that the verb specifies the (potential) affectedness of the container. We thus predict that relatively more container locatives for the holistic verb should be produced later in the sessions than earlier in the sessions. (In these preliminary experiments, we purposefully confounded greater exposure to holistic actions with exposure to aspectual cues.))

Two other major differences between the previous and next pair of studies concern their designs. In Experiments 3 and 4, we used a relatively powerful within-subject design, each subject being taught two verbs which varied according to whether the content or container was affected in a particular way. In the following two experiments, the identity of the (iterable) action across conditions forced us to use a less powerful between-subjects design, in which each subject is taught and tested on one verb meaning (holistic or partitive). These two types of design make up for each other's deficiencies: the more powerful within-subject design controls for individual differences, but must overcome large set effects; the less powerful between-subjects design avoids these set effects at the expense of factoring individual differences into the variation between verb meaning conditions. A second difference between these experiments and the preceding two involves the technique of eliciting locative utterances. In the following experiments we adopt the technique introduced, and used with considerable success, in Experiment 2.

Method

Subjects. Forty-eight children and sixteen adults, all native speakers of English living in the Boston area, participated in the study. The children fell into three age groups: sixteen between 3;5 and 4;10 (mean 4;0); sixteen between 5;0 and 6;11 (mean 5;7); and sixteen between 7;0 and 9;4 (mean 7;10). Seven children were replaced in the design (three for being uncooperative, inattentive, or shy; two for being unable to perform the production task; one for experimenter error; and one for receiving intervention deemed relevant to the performance on this task). The children were drawn from middle-class day-care and after-school programs in Cambridge, Newton, and Needham. The adults were MIT undergraduate and graduate students, ranging in age from 19 to 25 years, and were paid for their participation. (One adult was replaced in the design for misinterpreting the task as a request to assume the role, and capacities, of a first language learner!)

Materials. Two sets of materials were used in testing each subject. Each set included two types of objects and two containers. One set consisted of: 1/4" (diameter) wooden beads; 1/4" plastic eggs; a 3" x 8" wooden cart with six holes in its (single) surface (i.e., the cart was simply a board with four wheels attached); and a 4" square wooden cube with four holes on one of its sides. The second set consisted of: 3/4" glass marbles; 3/4" plastic balls; a 3" x 8" wooden bench with six holes in its horizontal surface; and a 3" x 24" wooden board with four holes in its surface. The objects and containers were interchangeable within each set, so that the objects of either type in a set (e.g., marbles or balls) could be inserted into (but never pushed through) the holes of either container in that set (e.g., the bench or the board). The number of holes on the surfaces varied from four or six, arranged in one row or two. Within subject, the same pairings of objects and containers were used throughout the session; across subjects in an

age group, the pairings of materials in one set were counterbalanced with the pairings of materials in the other set.

In addition, two (non-interchangeable) pairs of materials were used in the teaching phase of the experiment: 2" circular disks of styrofoam and a 6" x 10" aluminum muffin tray with 8 (cylindrical) cavities (arranged in 2 rows); 1" x 1" duplo rectangles and a 8" x 10" plastic candy mold with 12 rectangular indentations (arranged in 3 rows).

Two verb meanings were created using pairs of materials. In the partitive condition, the experimenter inserted one object into a hole in the container. In the holistic condition, the experimenter repeatedly inserted objects into the container, one at a time, until every hole in the container was covered. Each subject was taught and tested on one verb meaning. Across subjects in an age group, the partitive meaning was taught and tested as often as the holistic meaning. (Note: we made every effort to match the ages of the children between meaning conditions, so that (e.g.) the eight mid-aged children learning the holistic verb had the same mean age (± one month) as the eight mid-aged children learning the partitive verb.) In addition, we counterbalanced the four possible combinations of objects and containers with verb meaning so that each combination of object-container pairs (as well as each pair of materials) occurred as often in the partitive condition as it did in the holistic condition, across the subjects in an age group. Finally, corresponding to the one novel action taught to each subject was one stem, *keat*.

Procedure. The procedure consisted of five parts: first, the experimenter taught the subject a novel verb; second, the experimenter elicited locative utterances of the novel verb from the subject; third, the experimenter tested and re-trained the subject using aspectual cues; fourth, the experimenter again elicited locative utterances of the novel verb from the subject; fifth, the experimenter re-tested the subject using aspectual cues. Subjects were tested in a single (20-

minute) session by two experimenters (one eliciting responses; the other observing) in an area as free as possible from potential distractions. For the children, the novel verb was introduced as a puppet word by a puppet.

The experimenter began each session by introducing to the subject all of the materials used in the study. The experimenter also introduced (or had the puppet introduce) the phonetic form of the verb: "can you say *keat*? ... say *keat*." In teaching the novel verb to the subject, the experimenter performed the holistic or partitive action only once, using either the styrofoam piece(s) and muffin tray or the duplo piece(s) and candy mold. A minimum of linguistic information was provided during the performance of the action. In the partitive condition, the experimenter inserted (e.g.) a piece of styrofoam into a hole in the tray while saying, simultaneously, "I am keating." In the holistic condition, the experimenter inserted (e.g.) styrofoam pieces into the tray, one at a time, until all of the holes in the tray were covered; "I am keating" was uttered only once, but spaced over several iterations of the insertion action. The experimenter then asked the subject to perform the action once: "show me what keating is." The teaching was repeated for those subjects who failed to act out the verb meaning correctly, though this happened only rarely. The sparseness of the teaching phase, as well as the lack of any priming with English locatives as in Experiments 3 and 4, was designed to allow subjects to learn more about the meaning of the verb with more exposure to the action; in particular, it was thought that subjects might be able to revise their interpretations of the holistic verb in the light of the subsequent aspectual cues (and consequently, produce more container locatives later in the session).

After the teaching phase, the experimenter then began the first block of production trials. The procedure in these tests follows closely the protocol of the production task in Experiment 2. Besides posing either content-topic or container-topic queries, the experimenter presented

each subject with *two* potential materials for the non-topicalized participant in the action--that is, two types of objects if the container was topicalized and two containers if a type of object was topicalized. In Experiment 2, we found that this technique enabled us to elicit a greater proportion of full locatives (i.e., those with *into/onto*-phrases) by forcing subjects to identify which of the two non-topicalized materials was actually used in performing the action. We further motivated subjects to supply this information, as in Experiment 2, by having them describe the actions to a blindfolded puppet.

For example, after subjects were introduced to the blindfolded puppet Marty, and told that the purpose of the game was to tell Marty what was happening, the experimenter would use the following script to set up and pose a container-topic query (in the holistic condition): "Here is a board (topic)... I can have either some marbles (as experimenter points to a clear plastic bag of marbles)... or some balls (as experimenter points to a clear plastic bag of balls). Now watch this: I am keating (while experimenter performs the holistic action, as in the teaching phase, using the marbles and the board)... Tell Marty, using the word *keat*, what I did to the *board*." (Note that the order of presentation of the two potential materials was balanced within subject so that the chosen material was first as often as it was second.) In order to set up and pose a content-topic query (again, in the holistic condition), the experimenter would proceed as follows: (e.g.) "Here are some marbles (topic)... I can have either a board (as experimenter points to a board)... or a bench (as experimenter points to a bench). Now watch this: I am keating (while experimenter performs the holistic action, as in the teaching phase, using the marbles and the bench)... Tell Marty, using the word *keat*, what I did to the *marbles*."

The same scripts were used for the partitive action except that single objects, versus *sets* of objects, participated in the action, and thus were mentioned in the commentary and query: (e.g.) "Here is a ball" (when the object was topicalized); "I can have either a marble... or a

ball" (when the object was not topicalized); "Tell Mary, using the word *keat*, what I did to the *ball*?" (in the content-topic query). As in the previous experiments, we tested production at three levels of response, if necessary. (The order of choices in the tertiary query was balanced within subject.)

In the first block of production trials, the experimenter posed four queries to the subject. The order of topics for these queries was strictly alternating: either content-container-content-container or container-content-container-content. The order of query topics was balanced across the subjects in an age group. In addition, each performance of the novel action was performed with a new object and container, so that after four trials, each of the four objects and containers (excluding the teaching materials) had been used once. (The unchosen non-topicalized material on a given trial was always the remaining object or container in a set of four interchangeable materials.)

After the first block of production trials, the experimenter tested and re-trained the subject using aspectual cues. The testing consisted of the experimenter performing the partitive action (regardless of which verb meaning a subject had learned and had been production tested on), and asking, "am I done keating...did I keat?" The action was performed with the same pair of materials that was used in the initial training. We were most interested if subjects in the holistic condition answered *yes* to this question, indicating that they did not regard the novel verb as *necessarily* specifying the accomplishment of holistically affecting the container (cf. the endstate interpretation of *load*). Of course, we expected subjects in the partitive condition to always answer affirmatively.

Because answers to yes/no questions in child (language) research are notoriously non-demonstrative, we took a major function of this question to be to sensitize subjects to the endpoint of the novel action. Immediately following the question, subjects were "re-trained"

using aspectual cues (with the second pair of training materials, that they had not used before; the order of training materials was balanced across subjects in an age group). In the partitive condition, the experimenter inserted (e.g.) a duplo rectangle into a hole in the candy mold, saying simultaneously, "I am keating," and then afterwards, "I am done keating, I keated." In the holistic condition, the experimenter inserted duplo rectangles into the mold, one at a time, until all of the holes in the mold were covered; meanwhile, the experimenter said, "I am keating, but I am *not* done keating yet (after inserting an object, but before completing the action)... I am *not* done keating yet (after inserting another object, but before completing the action)... *now* I am done keating, I keated (after inserting an object, and completing the action)." The phrase *I am not done keating yet* was uttered at least twice for each performance of the holistic action. The experimenter then asked the subject to perform the action once (i.e., "show me what keating is"), during which time the experimenter interrupted subjects in the holistic condition to ask, "Are you done keating yet? ...(the experimenter supplied the correct answer if the subject failed to)."

Following the retraining, the experimenter began the second block of four production trials. The procedure here was the same as for the first block, except that aspectual cues were incorporated into the scripts for the holistic and partitive verbs. In the holistic condition, for example, the container-topic query would be set up as follows: (e.g.) "Here is a board (topic)... I can have either some marbles (as experimenter points to a clear plastic bag of marbles)... or some balls (as experimenter points to a clear plastic bag of balls). Now watch this: I am keating, but I am *not* done keating yet (after inserting a marble into the board, but before completing the action)... I am *not* done keating yet (after inserting another marble, but before completing the action)... *now* I am done keating, I keated (after inserting a marble, and completing the action)... Tell Marty, using the word *keat*, what I did to the *board*." In the partitive condition, the corresponding script would read: "Here is a board (topic)... I can have

either a marble (as experimenter points to a marble)... or a ball (as experimenter points to a ball). Now watch this: I am keating (while inserting a marble into the board)... I am done keating, I keated... Tell Marty, using the word *keat*, what I did to the *board*."

As in the first block, the order of topics for these queries was strictly alternating: either content-container-content-container or container-content-container-content. We counterbalanced the order of query topics for the first and second blocks, across subjects in an age group. In addition, each performance of the novel action in the second block was performed with a different object and container (but a pair that had been used once before in the first block). We coordinated the order of query topics with the order of material pairs so that each of the eight materials (excluding those used in teaching and retraining) was focused once and only once per session, and across subjects in an age group each material (in a given pairing) was focused as often in the partitive condition as it was in the holistic condition. As in Experiment 4, we chose this design in order to rule out the possibility that the focusing of different materials (with potentially different salience) could account for any observed effect of verb meaning.

In the final phase of the procedure, the experimenter re-tested the subject using aspectual cues. As in the initial testing, the experimenter performed the partitive action (with the same pair of materials that was used in the re-training), and then asked, "am I done keating...did I keat?" Again, we were most interested in the responses from subjects in the holistic condition. For example, would subjects in the holistic condition who had answered *yes* to this question the first time that it was asked (before the re-training) now answer *no*? If so, they may have learned that the novel verb must specify the accomplishment of holistically affecting the container (cf. the endstate interpretation of *fill*).

Scoring. The responses to the production tests were scored according to whether the direct object corresponded to the content or container in the performed action. (Acceptable forms also

included one passive (e.g., *the block was keated*) and two sentences in which the object (content) was encoded as an instrumental subject (e.g., *the bead keated the block*).) We used the conventions and protocol in the Scoring Section of Experiment 1 for determining whether or not the use of a pronoun (e.g., *keating it*) was ambiguous. Responses which were undecipherable or which included no specification of the direct object were coded as *other*. As in the previous experiments, responses were scored according to the level of response (1°, 2°, 3°) and according to whether oblique objects and/or particles were also uttered. We also recorded any spontaneous substitutions of English verbs for the novel ones being elicited, as well as any errors in, or unusual aspects of, a subject's performance in the experiment. The responses to the aspectual questions were scored as either *yes* or *no*.

Design. For the production task, we employed a 2 x 2 x 2 x 4 factorial design with the within-subject factors of Block Order (block 1 vs. block 2) and Query Topic (content vs. container), and the between-subjects factors of Meaning (partitive vs. holistic) and Age Group (3;5-4;10 vs. 5;0-6;11 vs. 7;0-9;4 vs. adult). The dependent variable was the proportion of trials in which either the content or container was encoded as direct object. For the aspectual comprehension task, we employed a 2 x 2 x 4 factorial design with the within-subject factor of Order (before retraining vs. after retraining) and the between-subjects factors of Meaning (partitive vs. holistic) and Age Group (3;5-4;10 vs. 5;0-6;11 vs. 7;0-9;4 vs. adult). The dependent variable was the proportion of trials in which the response was either *yes* or *no*.

Results and Discussion

We will address two questions. First, did children and adults produce relatively more content locatives for the verb in the partitive condition, and relatively more container locatives for the verb in the holistic condition? Second, were children and adults more willing to use the holistic verb in the container-locative form after hearing aspectual cues (i.e., in the second

block of production trials) than before hearing them (i.e., in the first block of production trials)? A final issue for discussion concerns the nature of the representation that language learners assign to holistically affected containers.

In Table 24 we present the proportion of trials in which subjects produced content and container locatives as a function of verb meaning, query topic, block order, and age group. As in Experiments 3 and 4, we used these means to derive a more useful dependent measure--the *preference score*: the proportion of trials in which a content locative is produced minus the proportion of trials in which a container locative is produced. Preference scores range from +1.0 (a strong preference for content locatives) to −1.0 (a strong preference for container locatives). Mean preference scores are listed in Table 25 as a function of verb meaning, query topic, block order and age group.

We performed an Analysis of Variance on the mean preference score, with the within-subject variables of Block Order and Query Topic and the between-subjects variables of Verb Meaning and Age Group. We found a main effect of verb meaning, indicating that subjects had a significantly higher preference score for the partitive verb (M = 0.62) than for the holistic verb (M = 0.24), $F(1, 56) = 4.36$, $p < .05$. We note that the effect was not found (across age groups) for responses to the primary query. Within age group, the effect was significant only for the mid-aged children (5;0-6;11, mean 5;7), for responses to the primary query (M_{hol} = -0.22, M_{par} = 0.52, $F(1, 14) = 6.48$, $p < .025$) as well as for responses to all levels of query (M_{hol} = -0.09, M_{par} = 0.75, $F(1, 14) = 4.91$, $p < .05$); however, the mean preference score was always lower in the holistic condition than in the partitive condition for each of the other groups, and the difference was significant (combining all levels of response) for the combined group of children (M_{hol} = 0.19, M_{par} = 0.65, $F(1, 46) = 4.63$, $p < .05$).

Table 24

Proportion of Trials in which Content and Container Locatives of Novel Verbs were Produced as a Function of Meaning, Query Topic, Block Order and Age Group

	AGE GROUP			
	3;5-4;10	5;0-6;11	7;0-9;4	Adult
PARTITIVE MEANING				
Content Locatives				
1st Block				
Content-Topic Query	0.88	0.88	0.81	0.94
Container-Topic Query	0.81	0.88	0.62	0.75
Mean	0.84	0.88	0.72	0.84
2nd Block				
Content-Topic Query	0.88	0.88	0.75	0.75
Container-Topic Query	0.88	0.88	0.75	0.69
Mean	0.88	0.88	0.75	0.72
Mean	0.86 (5/50/0)	0.88 (40/16/0)	0.73 (45/2/0)	0.78 (45/5/0)
Container Locatives				
1st Block				
Content-Topic Query	0.12	0.12	0.19	0.06
Container-Topic Query	0.19	0.12	0.38	0.25
Mean	0.16	0.12	0.28	0.16
2nd Block				
Content-Topic Query	0.12	0.12	0.25	0.25
Container-Topic Query	0.12	0.12	0.25	0.31
Mean	0.12	0.12	0.25	0.28
Mean	0.14 (4/5/0)	0.12 (7/1/0)	0.27 (13/4/0)	0.22 (11/3/0)

Note: The numerals in parentheses correspond to the frequencies of locatives produced at the 1°/2°/3° level of response.

Table 24 (Continued)

Proportion of Trials in which Content and Container Locatives of Novel Verbs were Produced as a Function of Meaning, Query Topic, Block Order and Age Group

	AGE GROUP			
	3;5-4;10	5;0-6;11	7;0-9;4	Adult
HOLISTIC MEANING				
Content Locatives				
1st Block				
Content-Topic Query	0.81	0.62	0.75	0.88
Container-Topic Query	0.56	0.38	0.69	0.50
Mean	0.69	0.50	0.72	0.69
2nd Block				
Content-Topic Query	0.81	0.56	0.62	0.81
Container-Topic Query	0.38	0.25	0.69	0.62
Mean	0.59	0.41	0.66	0.72
Mean	0.64 (25/16/0)	0.45 (16/13/0)	0.69 (42/2/0)	0.70 (45/0/0)
Container Locatives				
1st Block				
Content-Topic Query	0.19	0.38	0.25	0.12
Container-Topic Query	0.44	0.62	0.31	0.50
Mean	0.31	0.50	0.28	0.31
2nd Block				
Content-Topic Query	0.19	0.44	0.38	0.19
Container-Topic Query	0.62	0.75	0.31	0.38
Mean	0.41	0.59	0.34	0.28
Mean	0.36 (18/5/0)	0.55 (30/5/0)	0.31 (18/1/1)	0.30 (18/1/0)

Note: The numerals in parentheses correspond to the frequencies of locatives produced at the 1°/2°/3° level of response.

Table 25

Mean Preference Score for Novel Verbs as a Function of Meaning, Query Topic, Block Order, and Age Group

	AGE GROUP			
	3;5-4;10	5;0-6;11	7;0-9;4	Adult
PARTITIVE MEANING				
1st Block				
Content-Topic Query	0.75	0.75	0.62	0.88
Container-Topic Query	0.62	0.75	0.25	0.50
Mean	0.69	0.75	0.44	0.69
2nd Block				
Content-Topic Query	0.75	0.75	0.50	0.50
Container-Topic Query	0.75	0.75	0.50	0.38
Mean	0.75	0.75	0.50	0.44
Mean	0.72	0.75	0.47	0.56
HOLISTIC MEANING				
1st Block				
Content-Topic Query	0.62	0.25	0.50	0.75
Container-Topic Query	0.12	-0.25	0.38	0.00
Mean	0.38	0.00	0.44	0.38
2nd Block				
Content-Topic Query	0.62	0.12	0.25	0.62
Container-Topic Query	-0.25	-0.50	0.38	0.25
Mean	0.19	-0.19	0.31	0.44
Mean	0.28	-0.09	0.38	0.41

Note: Mean preference score was calculated by subtracting the mean proportion of trials in which container locatives were produced from the mean proportion of trials in which content locatives were produced.

We also found a significant main effect of query topic and a significant interaction between verb meaning and query topic. The main effect of query topic showed that subjects had a significantly higher preference score for responses to the content-topic query ($M = 0.58$) than for responses to the container-topic query ($M = 0.29$), $F(1, 56) = 16.55$, $p < .001$ (for primary responses, $M_{content\text{-}topic} = 0.40$, $M_{container\text{-}topic} = 0.16$, $F(1, 56) = 11.13$, $p < .005$.). This is the predicted result, replicating our previous findings, that subjects prefer to encode the topic of discourse (here, the query topic) as the direct object. The interaction of verb meaning and query topic indicates that the effect of query topic is significantly greater in the holistic condition ($M_{content\text{-}topic} = 0.47$, $M_{container\text{-}topic} = 0.02$) than in the partitive condition ($M_{content\text{-}topic} = 0.69$, $M_{container\text{-}topic} = 0.56$, $F(1, 56) = 5.33$, $p < .05$). This interaction was even stronger for responses to the primary query (in the holistic condition, $M_{content\text{-}topic} = 0.38$, $M_{container\text{-}topic} = -0.03$; in the partitive condition, $M_{content\text{-}topic} = 0.42$, $M_{container\text{-}topic} = 0.36$; $F(1, 56) = 5.99$, $p < .02$). It appears as though subjects selectively avoided the container-locative form of the partitive verb, despite the container-topic query. (We discuss this pattern of results below.) No other findings in the ANOVA were significant.

The main effect of verb meaning provides an answer to our first question: subjects did indeed produce relatively more content locatives for the verb in the partitive condition, and relatively more container locatives for the verb in the holistic condition. However, this effect appears to be small (albeit significant) compared to the effect of verb meaning in previous experiments. We found the effect of verb meaning in this experiment ($r = .27$) to be significantly smaller than the effect of verb meaning in Experiment 4 ($r = .80$; $Z = 4.54$, $p < .001$ (two-tailed)), and smaller, though not significantly, than the effect of verb meaning in Experiment 3 ($r = .51$; $Z = 1.58$, $p \approx .11$ (two-tailed)). In addition, the production of locatives appears to be skewed in favor of the content-locative form--with subjects showing an overall

preference for the content-locative form (M = 0.43), much as they did in Experiment 3 (M = 0.55). Indeed, the similarity goes further: in both experiments we find no preference for the syntax of one verb (i.e., the holistic verb and the stuffing/sagging verb), versus a preference for the content-locative form of the other verb (i.e., the partitive verb and the zig-zagging verb). A series of two-tailed t-tests against zero revealed that none of the age groups (including the combined group of children) had a preference score for the holistic verb that was significantly different from zero, whereas every group except the oldest children had a preference score for the partitive verb that was significantly greater than zero: M_{yng} = 0.72, $t(7) = 3.29$, $p < .02$; M_{mid} = 0.75, $t(7) = 3.00$, $p < .025$; M_{old} = 0.47, $t(7) = 1.69$, $p \approx .14$; M_{adt} = 0.56, $t(7) = 2.61$, $p < .05$; $M_{combined}$ = 0.65, $t(23) = 4.60$, $p < .001$.

The similarity of the present effect of verb meaning to that documented in Experiment 3, both in the size of the effect and in the (absolute) preference scores for the verbs in the two conditions, suggests that subjects may attribute both a manner and an endstate to the holistic verb, but only a manner to the partitive verb. This is consistent with our hypothesis concerning alternating verbs such as *load*, upon which the novel actions are modeled: we expect not only that subjects should be able to use affectedness to predict a difference between the syntax of holistic and partitive verbs, but also that they should selectively avoid uttering container locatives for the verb in the partitive condition because--in the context of the partitive action--the potential affectedness of the container specified by the verb has not been satisfied. Thus, subjects should avoid saying *you keated the board with the ball* in the partitive condition for the same reason that English speakers avoid saying *John loaded the gun with the bullet* (except in the special case when only one more bullet is needed to fill all of the chambers of the gun). By contrast, we have implicitly assumed that the affectedness of the content is not contingent on the extent of the action, which is consistent with our finding that subjects did not avoid uttering content locatives. Indeed, neither the children nor the adults preferred the container-locative

form of the holistic verb, despite the fact that *John loaded the gun with bullets* is more informative than, and presumably preferred to, *John loaded the bullets into the gun* in the case where the gun (as well as the set of bullets) is affected. It is possible that the Gricean edict to "be informative" does not apply in any straightforward way to the experimental context.

Turning to our second question--of whether children and adults were more willing to use the holistic verb in the container-locative form after hearing aspectual cues (i.e., in the second block of production trials) than before hearing them (i.e., in the first block of production trials)--we found that the preference score was indeed lower in the second block of holistic trials ($M = 0.19$) than in the first block ($M = 0.30$), but that the difference was not significant according to a two-tailed t-test ($t(28) = 1.16$, $p \approx .25$). For each of the child groups, however, we found that the preference score for the holistic verb was always (nonsignificantly) lower after retraining than before retraining, whereas the preference score for the partitive verb was never lower after retraining than before retraining. This interaction between verb meaning and block order approached significance for the combined child groups (in the holistic condition, $M_{1st\ block} = 0.27$, $M_{2nd\ block} = 0.10$; in the partitive condition, $M_{1st\ block} = 0.62$, $M_{2nd\ block} = 0.67$; $F(1, 46) = 3.08$, $p < .09$). We highlight this interaction because it argues against the notion that the retraining simply encouraged children to switch from their main response in the first block (i.e., from content to container locatives). Instead, this pattern of results suggests (weakly) that children were more confident in uttering the container-locative form of the holistic verb after hearing the aspectual cue (and with greater exposure to the holistic action) than before hearing the aspectual cue. On the other hand, we note that the adults did not perform in accordance with the hypothesis: they produced (nonsignificantly) *fewer* container locatives for the holistic verb after the retraining than before, and (nonsignificantly) *more* container locatives for the partitive verb after the retraining than before. Our evidence on the second question is therefore far from conclusive.

One final issue which we haven't addressed is whether subjects view the partitive/holistic distinction as a difference between verb meanings (cf., zig-zagging vs. stuffing/sagging) or as a difference between two situations to which one verb applies (cf., loading the bullet vs. loading the gun). (We have been equivocating between these interpretations--speaking of "verb meaning" on the one hand, and of "the extent of the action" on the other.) This is an important question because we have no assurance that subjects are even treating the novel verbs as akin to *load*: if they are, we have support not only for the ability of subjects to predict syntax on the basis of the extent of the action, but also for a particular analysis of (a class of) alternating verbs, according to which language learners must be able to separate the extent of the action from the meaning of the verb itself (see the General Discussion); if not, we still have support for the ability of subjects to predict syntax on the basis of meaning, but not necessarily for how this ability bears on the learning of alternating verbs. Although the between-subjects design of this (and the following) experiment does not allow us to disentangle this issue fully, we can make guarded use of the responses to the aspectual questions. In this task, the experimenter performed the partitive action, and then asked, "am I done keating...did I keat?" If a subject in the holistic condition answers *no* to this question, he or she probably regards the novel verb as *necessarily* specifying the accomplishment of holistically affecting the container (cf. *fill*). If a subject answers *yes*, he or she may *not* regard the novel verb as necessarily specifying the accomplishment of holistically affecting the container (cf. *load*).

In Table 26 we present the frequency of subjects who answered affirmatively to the aspectual questions, as a function of meaning, order (before retraining vs. after retraining), and age group. As expected, subjects in the partitive condition were unanimous (with the exception of one mid-aged child) in responding *yes* to both aspectual questions. Somewhat more surprising, however, is that ten subjects in the holistic condition (out of 32) answered *yes* to

Table 26

Frequency of Subjects Responding Affirmatively to the Aspectual Question as a Function of Meaning, Order, and Age Group

	AGE GROUP			
	3;5-4;10	5;0-6;11	7;0-9;4	Adult
PARTITIVE MEANING				
Before Retraining Only	0	0	0	0
After Retraining Only	0	1	0	0
Before and After Retraining	8	7	8	8
Total	8	8	8	8
HOLISTIC MEANING				
Before Retraining Only	3	0	1	1
After Retraining Only	0	1	0	0
Before and After Retraining	1	3	0	0
Total	4	4	1	1

Note: The Aspectual Question refers to the task in which the experimenter performed the partitive action, and then asked, "am I done keating...did I keat?" The "Before Retraining" Question was asked after the first block of production trials; the "After Retraining" Question was asked after the second block of production trials. In each age group there were 16 subjects.

one or both of the questions. Five of these subjects responded *yes* before retraining and *no* after retraining (three of the youngest children, one of the oldest children, and one adult), indicating that perhaps they had learned that the novel verb must specify the holistic affectedness of the container. Four responded *yes* before and after retraining (one of the youngest children and three of the mid-aged children), despite the retraining. Curiously, one mid-aged child responded *no* to the first question and *yes* to the second question, again despite the retraining inbetween. Although we cannot rule out the possibility that these ten subjects responded *yes* out of sheer compliance, another possibility is that these subjects may have taken the meaning of the novel verb to specify (at least initially) what we have called the "insertion action"--the (iterable) act of inserting one object into a hole (cf. *load*). Thus, they may have been able to factor apart the extent of the action from the meaning of the verb per se. (One problem with this conclusion is that the ten subjects in the holistic condition who responded *yes* to one or both of the aspectual questions had a mean preference score (M = 0.10) that was (nonsignificantly) lower than the remaining 22 subjects in the holistic condition (M = 0.31) who answered *no* to both aspectual questions. This is a slight embarrassment because we might expect subjects who treat the affectedness of the container as a necessity to produce more container locatives than those who treat the affectedness of the container as a contingency.)

A final bit of evidence comes from the English verbs which subjects spontaneously used during the production trials. Out of 512 trials (64 subjects x 8 trials), we found that subjects uttered the verb *put* on 94 (18%) occasions (plus the verbs *move* and *make*, uttered in one trial apiece). Of interest here is that the proportion of trials in which subjects uttered *put*, a verb similar in meaning to *insert*, is virtually identical for subjects in the partitive (49/256 or 0.19) and holistic (45/256 or 0.18) conditions. Although subjects may have uttered *put* because its

high frequency and general applicability (to many locative events), it is also possible that subjects viewed the meanings of the novel verbs similarly, in terms of the basic "insertion" relation, despite differences in the extent of the action. Furthermore, the preference scores on these trials are *both* elevated--in the holistic condition (M = 0.75) as well as the partitive condition (M = 0.83)--indicating that *put* also served as a syntactic model for these subjects. In the following experiment we shall explore the use of such models more systematically.

In summary, we found that subjects produced relatively more content locatives for the verb in the partitive condition than for the verb in the holistic condition. The relatively small size of this effect, and the finding that subjects selectively avoided uttering container locatives for the verb in the partitive condition, were taken as consistent with our hypothesis that the affectedness of the container, but not of the content, is contingent in some cases on the extent of the action. On the issue of aspectual cues, we found only weak evidence that children were more confident in uttering the container-locative form of the holistic verb after hearing the aspectual cue (and with greater exposure to the holistic action) than before hearing the aspectual cue. Finally, we presented evidence that some subjects factored apart the holistic extent of the action from the meaning of the verb itself, though we must regard the nature of the representation of holistically affected containers as still very much an open question. (We shall continue to regard the partitive/holistic distinction as one involving *meaning* at some level of lexical semantic representation, though perhaps not isolable to verb meaning.)

Experiment 6

One of the lessons we learned from Experiments 3-5 is that we couldn't completely control how a subject would interpret a novel verb. If we could, we might have expected the mean preference score for manner verbs always to be +1.00 and the mean preference score for endstate verbs always to be -1.00. The fact that subjects may vary in their construals of the same verb was made clear to us in the previous experiment, where we found that some of the subjects spontaneously used the English verb *put*, in lieu of the novel one, to describe the holistic action of covering all of the holes of a container. We also found that such subjects produced more content locatives than did the subjects in the holistic condition on average. Although we interpreted this spontaneous usage as evidence for the ability of subjects to make use of models in learning the syntax of new verbs, it is clear that a subject's spontaneous use of a model probably involves more than just the similarity of the model and the "target" (e.g., the holistic action); undoubtedly, the *accessibility* of the model to the subject is a crucial factor.

The purpose of the following study is to systematize the use of models, thereby demonstrating the syntactic consequences that may follow from how a novel verb is interpreted, by explicitly providing subjects with a choice between two familiar English verbs--*put* and *cover*. Our method is to prime subjects with both verbs in locative forms (e.g., *you put the piece of ribbon on the plate*;*you covered the cart with plastic*), ask them to choose one verb or the other as most similar to a new action, and finally elicit from them locative forms containing a novel verb for the new action. We predict that subjects who choose *put* as a model will produce relatively more content-locative forms of the novel verb and that subjects who choose *cover* as a model will produce relatively more container-locative forms of the novel verb.

If this prediction is borne out, what will we have shown? Let us, at the outset of this experiment, dispel the interpretation of such a finding as the trivial product of some process of "translation" or "analogy." There is no doubt that part of what we're studying here involves a conscious decision by the subject that the novel verb "is like putting" or "is like covering." The focus of our study, however, is on the lexical principles of mind that underlie the use of one model or another. By giving subjects a choice between two primes (rather than, say, arbitrarily separating subjects into two different priming conditions) we insure that the subject cannot avoid thinking about what the potential models and target *mean*.

Although we have endeavored to follow the format of Experiment 5 in designing this study, several major changes deserve mentioning. First, we have dispensed with the partitive action in this study, as the holistic action alone is an instance of both putting and covering. Furthermore, the holistic action itself has been changed to make it more similar to covering than the "insertion action" of Experiment 5; in the current variant, pieces of material are put onto a surface, one at a time, until the surface is completely covered. A final, major change is that we have added a control condition in which the holistic verb is taught and elicited without the benefit of primed models.

Method

Subjects. Seventy-two children and twenty-four adults, all native speakers of English living in the Boston area, participated in the study. Forty-eight children and sixteen adults participated in the main (model) condition; twenty-four children and eight adults participated in the control condition. The children in the main condition fell into three age groups corresponding closely to those of Experiment 5: sixteen between 3;7 and 4;10 (mean 4;5); sixteen between 5;0 and 6;10 (mean 5;6); and sixteen between 7;2 and 9;1 (mean 7;10). These age groups were closely matched (with those for the control condition: eight between 4;1 and

4;7 (mean 4;4); eight between 5;3 and 6;10 (mean 5;10); and eight between 7;0 and 8;4 (mean 7;9). Four children were replaced in the design for being unable to perform the production task. The children were drawn from middle-class day-care and after-school programs in Cambridge, Newton, Needham, and Weston. The adults (in the main and control conditions) were MIT undergraduate and graduate students, ranging in age from 16 to 26 years, and were paid for their participation.

Materials. As in Experiment 5, two sets of materials were used in eliciting novel forms from each subject. Each set included two types of objects and two containers. One set consisted of: 1 1/2" x 4" pieces of paper; 1 1/2" x 4" pieces of cloth; a 3" x 8" wooden cart (i.e., a board with four wheels attached); and a 4" x 9" wooden table. The second set consisted of: 2" x 4" pieces of felt; 2" x 4" pieces of ribbon; a 4" x 8" wooden bench; and a 6" x 8" wooden board. The objects and containers were interchangeable within each set, so that the objects of either type in a set (e.g., pieces of paper or pieces of cloth) could be placed onto the surface of either container in that set (e.g., the cart or the table) in such a way that the surface could be completely covered without objects overlapping each other or extending beyond the edges of the container. The number and configuration of objects on a surface was always constant within an object-container pair, but varied across pairs. Objects, either four or six in number, were placed either side-by-side in a row or end-to-end in two or three columns. Within subject, the same pairings of objects and containers were used throughout the session; across subjects in an age group, the pairings of materials in one set were counterbalanced with the pairings of materials in the other set.

One set of interchangeable materials was used for teaching the new action, eliciting the choice of model, and posing the aspectual question: 5" x 7" pieces of sponge; 5" x 7" pieces of styrofoam; a 10" x 14" rubber bath mat; and a 10" x 14" plastic tray. The number and

configuration of objects on a surface demarcated four quadrants on the surface. Within subject, the same pairings of objects and containers were used throughout the session; across subjects in an age group, the pairings of teaching materials were counterbalanced with the pairings of "production" materials above. In addition, a plastic plate and a large plastic sheet were used in the priming of locatives of *put* and *cover*, respectively.

One verb meaning was created using pairs of materials: the experimenter repeatedly placed objects onto the container (in the appropriate configuration), one at a time, until the surface of the container was completely covered without objects overlapping each other or extending beyond the edges of the container. Corresponding to the one novel action taught to each subject was one stem, *keat.*

Procedure. The procedure in the main condition was similar to that of the previous experiment, except for the addition of components before the first block of production trials (priming and eliciting models) and after the second block of production trials (eliciting models). In sum, the procedure consisted of eight parts: first, the experimenter primed locative forms of *put* and *cover*; second, the experimenter elicited the subject's choice of model for a new action; third, the experimenter taught the subject the novel verb corresponding to that new action; fourth, the experimenter elicited locative utterances of the novel verb from the subject; fifth, the experimenter tested and retrained the subject using aspectual cues; sixth, the experimenter again elicited locative utterances of the novel verb from the subject; seventh, the experimenter retested the subject using aspectual cues; and eighth, the experimenter again elicited the subject's choice of model for the action. The procedure for the control condition consisted of steps three-seven above (i.e., without priming or eliciting models). Subjects were tested in a single (25-minute) session by two experimenters (one eliciting responses; the other observing) in an area as free as

possible from potential distractions. For the children, the novel verb was introduced as a puppet word by a puppet.

The experimenter began each session by introducing to the subject all of the materials used in the study. The experimenter also introduced (or had the puppet introduce) the phonetic form of the verb: "can you say *keat*? ... say *keat*." The procedure for priming the English verbs *put* and *cover* in locative forms was similar to the priming (and pretesting) of *fill* and *pour* used in Experiments 3 and 4. In the present case, the priming consisted of asking subjects (unfocused) queries about actions of putting and covering, and subsequently giving subjects *feedback* on their responses--modeling the full content locative of *put* and the full container locative of *cover*. (We also performed an informal pretest by recording whether subjects encoded the content or container as the direct object of the verb.) For example, the experimenter would use the following script: "do you know what *putting* is?... watch this: (experimenter puts a piece of felt onto the plate)... can you tell me, using the word *put*, what I did?" In those cases where the query failed to elicit an unambiguous direct object, we followed up with the secondary prompt "putting what?" (A tertiary query was never needed.) Regardless of the subject's final response, the experimenter modeled a full content locative of *put* ("I'm putting the felt onto the plate"), and had the subject utter the sentence once, if not spontaneously before the feedback, then repeated after the feedback. The same procedure was followed for priming the container-locative form of *cover* (e.g., "I'm covering the cart with plastic"). (Note: we counted as permissible the utterance by subjects of container-locative forms of *cover* without oblique phrases, as in *you're covering the cart.*)

Put and *cover* were each primed four times in this fashion. The objects used for the priming of *put* were the four production objects (i.e., felt, ribbon, paper, and cloth); the container was always the plastic plate. The containers used for the priming of *cover* were the four production

containers (i.e., cart, table, bench, and board); the content was always the plastic sheet. (Using the materials in this way insured the subject's familiarity with all of the materials, without highlighting the salience of one material over another. Note: the order of materials used in the priming was balanced across subjects in an age group.) One verb was primed after the other, with the order of priming balanced across subjects within an age group.

After the priming (in the main condition), the experimenter introduced the subject to a new action, and elicited the subject's choice of model for the new action. The experimenter would say: (e.g.) "Now let's see something else. Watch this: (experimenter repeatedly places pieces of styrofoam onto the mat, one at a time in the appropriate configuration, until the mat is completely covered)... I'm going to show you again, but this time I'm going to ask you a question... (experimenter performs the action again)... using the word *put* or the word *cover*, can you tell me what I did?" We were most interested in which verb a subject would choose to describe the new action. In those cases where a subject responded with the utterance of *both* verbs, the experimenter asked, "if you had to choose just one, which would it be?" We also wanted to know which entity (object or container) the subject encoded as the direct object of the chosen verb. In those cases where the query failed to elicit (a verb plus) an unambiguous direct object, we followed up with the secondary prompt "putting what?" (if *put* was chosen by the subject) or "covering what?" (if *cover* was chosen by the subject). (A tertiary query was never needed.) After eliciting the subject's choice of model in this fashion, the experimenter switched to the other pair of teaching materials (e.g., the pieces of sponge and the tray), and repeated the question, "using the word *cover* or the word *put*, can you tell me what I did?" Besides changing the materials for the second question, the experimenter also switched the order of the verb choices (*cover* or *put*) in the query. The sequence of verb choices (*put* or *cover*, *cover* or *put* vs. *cover* or *put*, *put* or *cover*) was balanced across subjects in an age group.

There are several reasons why we used this procedure in order to elicit a subject's choice of model. Previous piloting had shown that the simpler question "is this [the new action] more like putting or more like covering?" was inadequate; subjects responded to this question by consciously comparing the three actions in all sorts of ways, most of which seemed irrelevant to the linking of verb meaning and verb syntax. Furthermore, the simpler forced choice provided us with no way of knowing which entity in the new action the subject would encode as the direct object of the chosen verb. The combined forced-choice & production question, on the other hand, appeared to be directly relevant to the issue of linking, while providing additional syntactic information about the models themselves.

After asking the model questions (in the main condition), the experimenter taught the subject the novel verb corresponding to the new action. The experimenter told the subjects: "puppets have a word for what I just did: *keat*... say *keat*." The experimenter then performed the new action again (switching back to the first set of teaching materials), saying "I am keating" once over the course of several iterations of the putting action. The experimenter then asked the subject to perform the action once: "show me what keating is." The teaching was repeated for those subjects who failed to act out the verb meaning correctly, though this happened only rarely. Children had no trouble in understanding that the novel label applied to the new action. As in Experiment 5, the sparseness of the teaching phase was designed to allow subjects to learn more about the meaning of the verb with more exposure to the action; in particular, it was thought that subjects might be able to revise their interpretations of the verb in the light of the subsequent aspectual cues (and consequently, produce more container locatives, or even switch their choice of models, later in the session).

Following the teaching of the novel verb, the experimenter began the first block of production trials. The procedure in the production blocks (and in the intervening aspectual

retraining) is virtually identical to the holistic condition in the previous experiment. In particular, the experimenter presented each subject with *two* potential materials for the non-topicalized participant in the action--that is, two types of objects if the container was topicalized and two containers if a type of object was topicalized. For example, after subjects were introduced to the blindfolded puppet Marty, and told that the purpose of the game was to tell Marty what was happening, the experimenter would use the following script to set up and pose a container-topic query: "Here is a board (topic)... I can have either some felt (as experimenter points to a clear plastic bag of felt pieces)... or some ribbon (as experimenter points to a clear plastic bag of ribbon pieces). Now watch this: I am keating (while experimenter performs the action, as in the teaching phase, using the ribbon and the board)... Tell Marty, using the word *keat*, what I did to the *board*." (Note that the order of presentation of the two potential materials was balanced within subject so that the chosen material was first as often as it was second.) In order to set up and pose a content-topic query, the experimenter would proceed as follows: (e.g.) "Here is some felt (topic)... I can have either a bench (as experimenter points to a bench)... or a board (as experimenter points to a board). Now watch this: I am keating (while experimenter performs the action, as in the teaching phase, using the bench and the felt)... Tell Marty, using the word *keat*, what I did to the *felt*." As in the previous experiments, we tested production at three levels of response, if necessary. (The order of choices in the tertiary query was balanced within subject.)

In the first block of production trials, the experimenter posed four queries to the subject. The order of topics for these queries was strictly alternating: either content-container-content-container or container-content-container-content. The order of query topics was balanced across the subjects in an age group. In addition, each performance of the novel action was performed with a new object and container, so that after four trials, each of the four objects and containers (excluding the teaching materials) had been used once. (The unchosen non-

topicalized material on a given trial was always the remaining object or container in a set of four interchangeable materials.)

After the first block of production trials, the experimenter tested and retrained the subject using aspectual cues. The testing consisted of the experimenter performing the partitive variant of the action--placing one object on a surface--and then asking, "am I done keating...did I keat?" The action was performed with the same pair of materials that was used in introducing the novel verb stem. As in the previous experiment, we were most interested if subjects answered *yes* to this question, indicating that they did not regard the novel verb as *necessarily* specifying the accomplishment of holistically affecting the container. Immediately following the question, subjects were "retrained" using aspectual cues (with the second pair of teaching materials). For example, the experimenter would place pieces of sponge onto the tray, one at a time, until the tray was covered; meanwhile, the experimenter would say, "I am keating, but I am *not* done keating yet (after placing an object on the surface, but before completing the action)... I am *not* done keating yet (after placing another object on the surface, but before completing the action)... *now* I am done keating, I keated (after placing yet another object on the surface, and completing the action)." The phrase *I am not done keating yet* was uttered at least twice for each performance of the action. The experimenter then asked the subject to perform the action once (i.e., "show me what keating is"), during which time the experimenter interrupted subjects to ask, "Are you done keating yet? ...(the experimenter supplied the correct answer if the subject failed to)."

The experimenter then began the second block of four production trials. The procedure here was the same as for the first block, except that aspectual cues were incorporated into the scripts. For example, the container-topic query would be set up as follows: (e.g.) "Here is a bo. d (topic)... I can have either some felt (as experimenter points to a clear plastic bag of felt

pieces)... or some ribbon (as experimenter points to a clear plastic bag of ribbon pieces). Now watch this: I am keating, but I am *not* done keating yet (after placing a piece of felt on the board, but before covering the board)... I am *not* done keating yet (after placing another piece of felt on the board, but before covering the board)... *now* I am done keating, I keated (after placing yet another piece of felt on the board, and covering it)... Tell Marty, using the word *keat*, what I did to the *board*."

As in the first block, the order of topics for these queries was strictly alternating: either content-container-content-container or container-content-container-content. We counterbalanced the order of query topics for the first and second blocks, across subjects in an age group. In addition, each performance of the novel action in the second block was performed with a different object and container (but a pair that had been used once before in the first block). As in Experiments 4 and 5, we coordinated the order of query topics with the order of material pairs so that each of the eight production materials was focused once and only once per session.

Following the second production block, the experimenter retested the subject using aspectual cues. As in the initial testing, the experimenter performed the partitive action (with the same pair of materials that was used in the retraining), and then asked, "am I done keating...did I keat?" Again, we were most interested if subjects answered *yes* to this question, or if they had answered *yes* to this question the first time that it was asked (before the re-training), and now answered *no*. (In the latter case, subjects may have learned that the novel verb must specify the accomplishment of holistically affecting the container (cf. *cover* or *fill*).)

At the close of the session (in the main condition), the experimenter again elicited the subject's choice of model for the action. The experimenter would say: (e.g.) "Now I have some new questions for you. Watch this: (experimenter repeatedly places pieces of styrofoam

onto the mat, one at a time in the appropriate configuration, until the mat is completely covered)... using the word *put* or the word *cover*, can you tell me what I did?" As in the first pair of model questions, we were most interested in which verb a subject would choose to describe the action, but also in whether or not a subject would switch models as the result of the aspectual cues and/or more exposure to the action. In those cases where a subject responded with the utterance of *both* verbs, the experimenter asked, "if you had to choose just one, which would it be?" We also recorded which entity (object or container) the subject encoded as the direct object of the chosen verb. In those cases where the query failed to elicit (a verb plus) an unambiguous direct object, we followed up with the secondary prompt "putting what?" (if *put* was chosen by the subject) or "covering what?" (if *cover* was chosen by the subject). (A tertiary query was never needed.)

After eliciting the subject's choice of model in this fashion, the experimenter switched to the other pair of teaching materials (e.g., the pieces of sponge and the tray), and repeated the question, "using the word *cover* or the word *put*, can you tell me what I did?" The use of teaching materials was balanced within subject so that each material was used equally often in the model questions (each pair used twice), as well as in teaching/retraining (each pair used once) and in the aspectual questions (each pair used once). Across subjects in an age group, the order and pairing of materials in the teaching set were counterbalanced. Besides changing the materials for the second question, the experimenter also switched the order of the verb choices (*cover* or *put*) in the query. Across subjects in an age group, the sequence of verb choices in the first pair of model questions was counterbalanced with the sequence of choices in the second pair of model questions.

Scoring. The responses to the production tests were scored according to whether the direct object corresponded to the content or container in the performed action. (Acceptable forms also

included one unaccusative intransitive (e.g., *the felt keated across the board*) and fifteen sentences (all uttered by adults) in which the object (content) was encoded as an instrumental subject (e.g., *she used the felt to keat the block*).) We used the conventions and protocol in the Scoring Section of Experiment 1 for determining whether or not the use of a pronoun (e.g., *keating it*) was ambiguous. Responses which were undecipherable or which included no specification of the direct object were coded as *other*. As in the previous experiments, responses were scored according to the level of response (1°, 2°, 3°) and according to whether oblique objects and/or particles were also uttered. We also recorded any spontaneous substitutions of English verbs for the novel ones being elicited, as well as any errors in, or unusual aspects of, a subject's performance in the experiment. The responses to the model questions were scored according to the verb chosen, *put* or *cover*, and according to whether the direct object corresponded to the content or container in the performed action (using the scoring procedure above). The responses to the aspectual questions were scored as either *yes* or *no*.

Design. For the production task, we employed a 2 x 2 x 2 x 4 factorial design with the within-subject factors of Block Order (block 1 vs. block 2) and Query Topic (content vs. container), and the between-subjects factors of Model Condition (main vs. control) and Age Group (youngest vs. mid-aged vs. oldest vs. adult). The dependent variable was the proportion of trials in which either the content or container was encoded as direct object. For the model elicitation task, we employed a 2 x 4 factorial design with the within-subject factor of Order (before production vs. after production) and the between-subjects factor of Age Group (youngest vs. mid-aged vs. oldest vs. adult). The dependent variables were the choice of model (*put* or *cover*) and the proportion of trials in which either the content or container was encoded as direct object. For the aspectual comprehension task, we employed a 2 x 2 x 4 factorial design with the within-subject factor of Order (before retraining vs. after retraining) and the between-subjects factors of Model Condition (main vs. control) and Age Group

(youngest vs. mid-aged vs. oldest vs. adult). The dependent variable was the proportion of trials in which the response was either *yes* or *no*.

Results and Discussion

We will address four questions. First, were children and adults willing to produce both content and container locatives for the holistic verb (i.e., showing no absolute preference for either form), as was the case in Experiment 5? Second, did subjects who chose *put* as a model produce relatively more content-locative forms of the novel verb, and did those who chose *cover* as a model produce relatively more container-locative forms of the novel verb? Third, if subjects failed to choose *put* or *cover* as a model, what were the consequences for the production of the novel verb. Fourth, what do the responses to the aspectual questions tell us about how children and adults interpreted the holistic extent of the novel action? In this last regard, we will also look at whether the subjects' choice of model had any effect upon their responses to the aspectual questions, and whether the subjects' responses to the aspectual questions, and/or more exposure to the holistic action, had any consequences for the production of locatives with the novel verb.

In Table 27 we present the proportion of trials in which subjects produced content and container locatives as a function of model condition (main vs. control), query topic, block order, and age group. As in Experiments 3-5, we used these means to derive a more useful dependent measure--the *preference score*: the proportion of trials in which a content locative is produced minus the proportion of trials in which a container locative is produced. Preference scores range from +1.0 (a strong preference for content locatives) to –1.0 (a strong preference for container locatives). Mean preference scores are listed in Table 28 as a function of model condition, query topic, block order, and age group.

Table 27

Proportion of Trials in which Content and Container Locatives of the Novel Verb were Produced as a Function of Model Condition, Query Topic, Block Order and Age Group

	AGE GROUP			
MAIN CONDITION	3;7-4;10	5;0 6;10	7;2-9;1	Adult
Content Locatives				
1st Block				
Content-Topic Query	0.28	0.28	0.53	0.25
Container-Topic Query	0.12	0.12	0.31	0.09
Mean	0.20	0.20	0.42	0.17
2nd Block				
Content-Topic Query	0.31	0.22	0.56	0.28
Container-Topic Query	0.09	0.12	0.44	0.09
Mean	0.20	0.17	0.50	0.19
Mean	0.20 (22/4/0)	0.19 (18/5/1)	0.46 (56/3/0)	0.18 (23/0/0)
Container Locatives				
1st Block				
Content-Topic Query	0.72	0.69	0.47	0.75
Container-Topic Query	0.88	0.88	0.69	0.91
Mean	0.80	0.78	0.58	0.83
2nd Block				
Content-Topic Query	0.66	0.78	0.44	0.72
Container-Topic Query	0.91	0.88	0.56	0.91
Mean	0.78	0.83	0.50	0.81
Mean	0.79 (77/24/0)	0.80 (76/26/1)	0.54 (65/4/0)	0.82 (105/0/0)

Note: The numerals in parentheses correspond to the frequencies of locatives produced at the 1°/2°/3° level of response.

Table 27 (Continued)

Proportion of Trials in which Content and Container Locatives of the Novel Verb were Produced as a Function of Model Condition, Query Topic, Block Order and Age Group

	AGE GROUP			
	4;1-4;7	5;3-6;10	7;0-8;4	Adult
CONTROL CONDITION				
Content Locatives				
1st Block				
Content-Topic Query	0.62	0.50	0.69	0.19
Container-Topic Query	0.50	0.50	0.62	0.00
Mean	0.56	0.50	0.66	0.09
2nd Block				
Content-Topic Query	0.44	0.50	0.50	0.19
Container-Topic Query	0.50	0.50	0.62	0.12
Mean	0.47	0.50	0.56	0.16
Mean	0.52 (19/14/0)	0.50 (12/19/1)	0.61 (29/10/0)	0.12 (8/0/0)
Container Locatives				
1st Block				
Content-Topic Query	0.38	0.50	0.31	0.81
Container-Topic Query	0.50	0.50	0.38	1.00
Mean	0.44	0.50	0.34	0.91
2nd Block				
Content-Topic Query	0.56	0.50	0.50	0.81
Container-Topic Query	0.50	0.50	0.38	0.88
Mean	0.53	0.50	0.44	0.84
Mean	0.48 (17/14/0)	0.50 (21/11/0)	0.39 (16/9/0)	0.88 (56/0/0)

Note: The numerals in parentheses correspond to the frequencies of locatives produced at the 1°/2°/3° level of response.

Table 28

Mean Preference Score for the Novel Verb as a Function of Model Condition, Query Topic, Block Order, and Age Group

MAIN CONDITION

	AGE GROUP			
	3;7-4;10	5;0-6;10	7;2-9;1	Adult
1st Block				
Content-Topic Query	-0.44	-0.41	0.06	-0.50
Container-Topic Query	-0.75	-0.75	-0.38	-0.81
Mean	-0.59	-0.58	-0.16	-0.66
2nd Block				
Content-Topic Query	-0.34	-0.56	0.12	-0.44
Container-Topic Query	-0.81	-0.75	-0.12	-0.81
Mean	-0.58	-0.66	0.00	-0.62
Mean	-0.59	-0.62	-0.08	-0.64

CONTROL CONDITION

	AGE GROUP			
	4;1-4;7	5;3-6;10	7;0-8;4	Adult
1st Block				
Content-Topic Query	0.25	0.00	0.38	-0.62
Container-Topic Query	0.00	0.00	0.25	-1.00
Mean	0.12	0.00	0.31	-0.81
2nd Block				
Content-Topic Query	-0.12	0.00	0.00	-0.62
Container-Topic Query	0.00	0.00	0.25	-0.75
Mean	-0.06	0.00	0.12	-0.69
Mean	0.03	0.00	0.22	-0.75

Note: Mean preference score was calculated by subtracting the mean proportion of trials in which container locatives were produced from the mean proportion of trials in which content locatives were produced.

We performed an Analysis of Variance on the mean preference score, with the within-subject variables of Block Order and Query Topic and the between-subjects variables of Model Condition and Age Group . We found significant main effects of model condition, age group, and query topic, and a significant interaction between model condition and query topic. The main effect of model condition indicated that the preference score was significantly lower for the subjects in the main condition (M = -0.48) than for those in the control condition (M = -0.12), $F(1, 88) = 4.75$, $p < .05$. This effect was marginally significant for responses to the primary query: M_{main} = -0.40, $M_{control}$ = -0.16, $F(1, 88) = 2.96$, $p < .09$. Within age groups, we found marginally significant effects for the youngest (M_{main} = -0.59, $M_{control}$ = 0.03, $F(1, 22) = 3.95$, $p < .06$) and mid-aged children (M_{main} = -0.62, $M_{control}$ = 0.00, $F(1, 22) = 3.20$, $p < .09$), but not for the oldest children (M_{main} = -0.08, $M_{control}$ = 0.22, $p > .25$) or adults (M_{main} = -0.64, $M_{control}$ = -0.75, $p > .25$). For the combined group of children, the effect was significant at the primary level of response (M_{main} = -0.32, $M_{control}$ = 0.03, $F(1, 70) = 4.70$, $p < .05$) as well as for all levels of response (M_{main} = -0.43, $M_{control}$ = 0.08, $F(1, 70) = 6.43$, $p < .02$).

The main effect of query topic indicates that subjects had a significantly higher preference score for responses to the content-topic query (M = -0.24) than for responses to the container-topic query (M = -0.48), $F(1, 88) = 17.18$, $p < .001$ (for primary responses, $M_{content\text{-}topic}$ = -0.21, $M_{container\text{-}topic}$ = -0.43, $F(1, 88) = 14.18$, $p < .001$.) This finding shows, once again, the predicted influence of discourse topic on choice of direct object. We also found a significant interaction between model condition and query topic, according to which the effect of query topic was significantly greater in the main condition ($M_{content\text{-}topic}$ = -0.31, $M_{container\text{-}topic}$ = -0.65) than in the control condition ($M_{content\text{-}topic}$ = -0.09, $M_{container\text{-}topic}$ = -0.16, $F(1, 88) = 4.76$, $p < .05$). This interaction was marginally significant for

responses to the primary query (in the main condition, $M_{content\text{-}topic}$ = -0.24, $M_{container\text{-}topic}$ = -0.55; in the control condition, $M_{content\text{-}topic}$ = -0.13, $M_{container\text{-}topic}$ = -0.20; $F(1, 88) = 3.75, p < .06$). It seems likely that the accessibility of the models to subjects in the main condition gave them concrete examples of how to interpret the focused queries.

The main effect of age group indicates that the preference score was significantly different for different age groups: M_{yng} = -0.38, M_{mid} = -0.41, M_{old} = 0.02, M_{adt} = -0.68, $F(3, 88) = 3.50, p < .02$ (for primary responses, M_{yng} = -0.28, M_{mid} = -0.35, M_{old} = 0.02, M_{adt} = -0.68, $F(3, 88) = 4.99, p < .005$). Roughly speaking, these results show that the oldest children had a higher mean preference score than the other groups. (We shall offer an explanation below involving the ability of subjects to use models in learning the syntax of novel verbs.) No other findings in the ANOVA were significant.

The main effect of model condition bears on our first question--of whether subjects were willing to produce both content and container locatives for the holistic verb. In Experiment 5, we found that none of the age groups had a preference score for the holistic verb that was significantly different from zero. In the present experiment, we find much the same result for the child (and combined child) groups in the control condition (M_{yng} = 0.03, $p > .25$; M_{mid} = 0.00, $p > .25$; M_{old} = 0.22, $p > .25$; $M_{combined}$ = 0.08, $p > .25$). In the main condition, by contrast, the preference scores were negative for every age group, and significantly less than zero for the younger (and combined) children: M_{yng} = -0.59, $t(15) = -3.78, p < .005$; M_{mid} = -0.62, $t(15) = -3.91, p < .002$; M_{old} = -0.08, $p > .25$; $M_{combined}$ = -0.43, $t(47) = -4.01, p < .001$. A further result is that while the (younger) children show an absolute preference for the container-locative form of the verb only in the main condition, the adults consistently preferred the container-locative form in both the main condition (M = -0.64, $t(15) = -4.14, p < .001$) and the control condition (M = -0.75, $t(7) = -3.97, p < .01$). The significant discrepancy between

the adult performance here (in the control condition) and in Experiment 5 ($M = 0.41$, $t(14) = 3.57$, $p < .005$)) suggests that the affectedness of the container was much more obvious in the "covering" action than in the holistic "pegging" action of Experiment 5. This makes sense if we consider the great similarity of the "covering" action, in which an entire surface of a container is occluded by content, to the actions lexicalized by common locative verbs in the language (e.g., *cover*, *spread*, and *blanket*, among others); by contrast, the holistic verb of Experiment 5 is similar to fewer verbs, and verbs of lower frequency--such as *load*, *stud*, and *sow*.

We can also make sense of the differences between the results for the younger children and adults if we assume that common locative verbs such as *cover* may be good models for the novel verb in this experiment only if such models are *accessible* to the subjects. In particular, we suggest that the priming and eliciting of models made the verb *cover* accessible to the younger children in the main condition. (Based on the verbs uttered spontaneously in Experiment 5, we assume that *put* was already accessible.) Without a particularly appropriate model, children in the control condition, as in Experiment 5, may have been equally likely to attribute a manner to the holistic verb (e.g., placing the objects on the surface in a particular configuration) as an endstate. According to the linking hypothesis, then, these children would show no preference for either form, as we have shown. In contrast, the younger children in the main condition (who were provided with access to *cover*) and the adults in the main and control conditions (who already have access to *cover* by virtue of being mature speakers of English) may have attributed primarily an endstate to the novel verb. These subjects would then be expected to produce more container locatives than content locatives, accounting for the results above. We note that the oldest children had no preference for either locative form; their performance is a bit mysterious on this account. We shall clarify their performance below.

If making models accessible to younger children has such an impact on their syntactic preferences, we might expect the actual *choice* of models to have an even larger, more clear-cut effect--potentially applying to older children and adults as well. In particular, did subjects who chose *put* as a model produce relatively more content-locative forms of the novel verb, and did those who chose *cover* as a model produce relatively more container-locative forms of the novel verb? To answer this question, we classified subjects according to their pattern of responses to the model questions. In general, we found that there was agreement between the utterance of a model's name and its syntactic form, so that *put* was usually (96%) uttered in a content-locative form and *cover* was usually (98%) uttered in a container-locative form. We observed utterances such as "putting a mat" and "you covered the styrofoam on the tray" in four instances each. These deviations from standard usage may have been induced by the demands of the model question (subjects may have felt compelled to use both models, despite an underlying preference), or they may reflect genuine lexical knowledge. In any case, when confronted with a conflict between the observed syntactic form and the one predicted on the basis of standard usage, we regarded the observed syntax as a more reliable metric of how a child would use a particular model.

In Table 29 we list the frequencies of subjects, by age group, falling into nine categories according to their pattern of responses to the two pairs of model questions (one pair before the production task; one pair after the production task). Ignoring the variable of order for the present, the nine categories may be collapsed into three types, involving the elicitation of: more container-locative models (3 or 4) than content-locative models; more content-locative models (3 or 4) than container-locative models; and equal numbers (2 each) of content- and container-locative models. Let us consider subjects performing according to the first type of pattern to have *chosen* the *cover* model, and those performing according to the second type of pattern to

Table 29

Frequency of Subjects Classified According to Type and Category of Response in the Model Elicitation Task and Age Group

	AGE GROUP			
	3;7-4;10	5;0-6;10	7;2-9;1	Adult
TYPE OF RESPONSE CATEGORY (1st Model Pair; 2nd Model Pair)				
Subjects Choosing *Cover*				
2 Container Models; 2 Container Models	7	5	2	10
2 Container Models; Each Model	1	1	1	1
Each Model; 2 Container Models	1	2	1	1
Total	9	8	4	12
Subjects Choosing *Put*				
2 Content Models; 2 Content Models	2	2	1	1
2 Content Models; Each Model	1	0	1	0
Each Model; 2 Content Models	1	2	2	2
Total	4	4	4	3
Subjects Splitting their Responses				
2 Container Models; 2 Content Models	0	1	0	0
2 Content Models; 2 Container Models	2	0	0	0
Each Model; Each Model	1	3	8	1
Total	3	4	8	1

Note: Four models were elicited from each subject in the Model Elicitation Task: The first pair was elicited before the novel-verb production trials; the second pair was elicited after the novel-verb production trials. The term *Each Model* signifies that one content-locative model and one container-locative model were elicited as a pair.

Table 30

Mean Preference Score for the Novel Verb as a Function of Type of Response in the Model Elicitation Task and Age Group

	AGE GROUP			
	3;7-4;10	5;0-6;10	7;2-9;1	Adult
TYPE OF RESPONSE				
Subjects Choosing *Cover*	-0.79	-0.97	-0.62	-0.79
Subjects Choosing *Put*	0.06	-0.31	-0.19	0.08
Subjects Splitting their Responses	-0.83	-0.22	0.25	-1.00

Note: Mean preference score was calculated by subtracting the mean proportion of trials in which container locatives were produced from the mean proportion of trials in which content locatives were produced.

have *chosen* the *put* model. Using these criteria, we performed a series of two-tailed *t*-tests on the difference in mean preference score (in the production task) between subjects choosing *put* as a model and subjects choosing *cover* as a model. The mean preference scores for subjects in an age group choosing *put* or *cover* as a model are listed in Table 30. Our findings provide strong support for the hypothesis that subjects have the ability to make use of models in learning the syntax of new verbs. Across age groups, we found that subjects who chose *cover* (M = -0.81) had a significantly lower preference score than those who chose *put* (M = -0.10), $t(46) = 4.18$, $p < .001$. This result was also highly significant for responses to the primary query: M_{cover} = -0.68, M_{put} = -0.09, $t(46) = 3.61$, $p < .001$. Within age groups, we found that the mean preference score was always lower for those subjects who had chosen *cover* as their model, significantly so for every group except the oldest children: for the youngest children, M_{cover} = -0.79, M_{put} = 0.06, $t(11) = 2.57$, $p < .05$; for the mid-aged children, M_{cover} = -0.97, M_{put} = -0.31, $t(10) = 2.27$, $p < .05$; for the oldest children, M_{cover} = -0.62, M_{put} = -0.19, $t(6) = 0.78$, $p > .25$; for the combined children, M_{cover} = -0.83, M_{put} = -0.15, $t(31) = 3.32$, $p < .005$; for the adults, M_{cover} = -0.79, M_{put} = 0.08, $t(13) = 2.52$, $p < .05$.

What about those subjects who failed to choose *put* or *cover* as a model? To answer this question, we focused our attention on subjects who *split* their responses in the model elicitation task, selecting equal numbers of the content- and container-locative models (see Table 30). We performed a series of two-tailed *t*-tests on the difference in mean preference score between subjects who split their responses and subjects who chose *put* or *cover* as a model. Across age groups, we found that the mean preference score for subjects who split their responses (M = -0.15) was significantly higher than the mean preference score for subjects who chose *cover* as a model (M = -0.81, $F(1, 47) = 3.66$, $p < .001$; for primary responses, M_{split} = -0.10, M_{cover} = -0.68, $F(1, 47) = 3.28$, $p < .002$), but not significantly different than the mean

preference score for subjects who chose *put* as a model (M = -0.10, $p > .25$). Within age groups, however, the pattern of results was somewhat variable. For subjects who split their responses versus those who chose *cover*, significant differences were found: for the mid-aged children (M_{split} = -0.22, M_{cover} = -0.97, $t(10) = 2.87$, $p < .02$); marginally for the older children (M_{split} = 0.25, M_{cover} = -0.62, $t(10) = 1.82$, $p < .10$); and for the combined children (M_{split} = -0.09, M_{cover} = -0.83, $t(34) = 3.63$, $p < .001$). For subjects who split their responses versus those who chose *put*, significant differences were found only for the adults (M_{split} = -1.00, M_{put} = 0.08, $t(2) = 6.50$, $p < .05$).

In general, we take these results to complement our earlier conclusion that models must be accessible in order to be useful; they must also be chosen (esp. in the case of *cover*) to be useful. In the experimental setting, of course, the choice of a model must take the form of an overt response; the subject must make the conscious decision that the novel action "is like putting" or "is like covering." There should be no doubt, however, that the influences of a learner's previously mastered lexical knowledge on the learning of new verbs may be unconscious and quite subtle. For this reason, in fact, it is difficult to say why we found no significant difference in mean preference score between subjects who split their responses and those who chose *put*. Subjects may have brought to the task more lexical knowledge of the content-locative form than of the container-locative form, regardless of their explicit responses to the model question. On the other hand, the small number of subjects who split their responses (16) or chose *put* (15; cf. 33 for *cover*) leads one to suspect the possibility of sampling error.

Assuming that subjects of all ages can use models in learning the syntax of new verbs, we can now understand our earlier findings that the oldest children had no absolute preference for either locative form, and a higher mean preference score than the other groups: more of these

subjects split their responses to the model question, and fewer made use of either model in learning the syntax of the novel verb. In particular, an examination of their responses to the model questions reveals that half of the oldest children (8) failed to choose *put* or *cover* as a model; this is as many as for the other groups combined (three of the youngest children, four of the mid-aged children, and one of the adults split their model responses). Crucially, the mean preference score for the oldest children who split their model responses was 0.25, marginally higher than the preference score for the four oldest children who chose *cover* as a model (M = -0.62, $t(10) = 1.82$, $p < .10$). Thus, the performance of the oldest children in this experiment is less of a mystery if we take into consideration their choice of models, or lack thereof.

A final topic that we will consider is the aspectual comprehension task: what can it tell us about how children and adults interpreted the holistic extent of the novel action? In Table 31 we present the frequency of subjects who answered *yes* to the aspectual questions, as a function of model condition, order (before retraining vs. after retraining), and age group. As in Experiment 5, we again found that some of the younger children were willing to consider the transfer of a single object as an instance of the novel action. In particular, 13 of the youngest children (out of 24) answered *yes* to the first question (11) or to both questions (2); seven of the mid-aged children (out of 24) answered *yes* to the first question (6) or to both questions (1). In contrast, only one of the mid-aged children, and none of the adults, responded *yes* to an aspectual query. Our (tentative) interpretation of these results is that the adults and oldest children regarded the holistic action as *necessarily* specifying the accomplishment of holistically affecting the container (cf. *cover* or *fill*), whereas at least some of the younger children (initially) regarded the holism of the action as a contingent property of the verb (cf. the endstate interpretation of *load*) or even as an "accident" of circumstance (cf. *put*). The results also show

Table 31

Frequency of Subjects Responding Affirmatively to the Aspectual Question as a Function of Model Condition, Order, and Age Group

MAIN CONDITION

	AGE GROUP			
	3;7-4;10	5;0-6;10	7;2-9;1	Adult
Before Retraining Only	7	3	1	0
After Retraining Only	0	0	0	0
Before and After Retraining	2	1	0	0
Total	9	4	1	0

CONTROL CONDITION

	AGE GROUP			
	4;1-4;7	5;3-6;10	7;0-8;4	Adult
Before Retraining Only	3	3	0	0
After Retraining Only	0	0	0	0
Before and After Retraining	1	0	0	0
Total	4	3	0	0

Note: The Aspectual Question refers to the task in which the experimenter performed the partitive action, and then asked, "am I done keating...did I keat?" The "Before Retraining" Question was asked after the first block of production trials; the "After Retraining" Question was asked after the second block of production trials. In each age group in the main condition, there were 16 subjects; in each age group in the control condition, there were 8 subjects.

the efficacy of the retraining; of the 21 children (including the one mid-aged child) who initially responded *yes* to the aspectual question, all but three later responded *no*.

In the previous experiment, it was impossible to judge whether the subjects who responded *yes* did so out of sheer compliance. In the model condition of this experiment, though, we have an independent source of information about how children and adults interpreted the holistic extent of the novel action; in particular, we can look at whether the younger children's initial choice of a model had any systematic effect upon their responses to the initial aspectual question (i.e., before retraining; given the efficacy of the retraining, we confined this comparison to the initial responses in either task). Our question is this: was it the case that younger children who initially chose *cover* as a model tended to respond *no* to the first aspectual query, whereas those who initially chose *put* as a model tended to respond *yes* to the first aspectual query? In order to answer this question, we constructed a 2 x 2 contingency table in which younger children were scored as responding with either two content-locative or two container-locative models and either a *yes* or *no*. We found that three children initially chose *put* and responded *yes*; four children initially chose *put* and responded *no*; five children initially chose *cover* and responded *yes*; and ten children initially chose *cover* and responded *no*. According to a Fisher Exact Test, there was no significant association between the initial responses in these two comprehension tasks. (Furthermore, in a test of *all* of the model responses, we found the same ratio of subjects choosing *put* to subjects choosing *cover* for both the "*yes* subjects" and the "*no* subjects.") Of course, this result does not prove that our younger subjects responded *yes* out of the desire to comply with the experimenter, but it strongly raises that possibility.

A final issue concerning the aspectual questions involves whether the subjects' responses to the aspectual questions, and/or more exposure to the holistic action, had any consequences

for the production of locatives with the novel verb. In Experiment 5 we found that subjects produced more container locatives after retraining than before, but that this difference was not significant. In addition, we discovered that the ten subjects in the holistic condition who responded *yes* to one or both of the aspectual questions had a mean preference score (M = 0.10) that was (nonsignificantly) lower than the remaining 22 subjects in the holistic condition (M = 0.31) who answered *no* to both aspectual questions. As we noted there, this is a problem because subjects who treat the affectedness of the container as a necessity (i.e., answering *no*) should produce more container locatives than those who treat the affectedness of the container as a contingency (i.e., answering *yes*). In the present experiment, we found much the same results: there was no effect of block order ($M_{1st\ block}$ = -0.36, $M_{2nd\ block}$ = -0.36), and the 21 subjects who responded *yes* to the aspectual question did not have a higher mean preference score (M = -0.38) than the 75 subjects who responded *no* (M = -0.36). The combined results of Experiments 5 and 6 suggest that overt aspectual cues, at least as we have envisioned them here, play a minor role in verb learning compared with that of a semantically (and by hypothesis, syntactically) similar model. In our view, however, the issue of how a child perceives the distribution of an action over time, and the relation of that perception to verb learning, demands much more study.

In summary, we found that children in the control condition were willing to produce both content and container locatives, as in experiment 5, but that children in the model condition and adults in either condition consistently preferred the container-locative form of the (holistic) verb. To account for this pattern of preferences across ages and conditions, we suggested that the linking hypothesis must be understood in the context of what models are accessible to language learners and speakers. In particular, we suggested that the priming and eliciting of models made the verb *cover* accessible to the younger children in the main condition, whereas adults (in either condition) already had access to *cover* by virtue of being mature speakers of

English. Beyond accessibility, however, we found that the actual *choice* of models had an even larger, more clear-cut effect--applying directly to the oldest children and adults as well to the younger children. Across age groups, we found that subjects who chose *cover* as a model had a significantly lower preference score than those who chose *put* as a model. Furthermore, for subjects who failed to choose *put* or *cover* as a model, especially the oldest children, we found less systematic consequences for the production of locative forms of the novel verb. We took these findings to provide strong support for the hypothesis that subjects of all ages have the ability to make use of models in learning the syntax of new verbs. Finally, on the issue of the aspectual cues and questions, we again found--as in Experiment 5--that some of the younger children responded *yes* to the aspectual questions. Although these younger children may have regarded the holism of the action as a contingent property of the verb (cf. the endstate interpretation of *load*) or even as an accident of how the actions were performed (cf. *put*), we failed to find support for this conclusion either from tests of association between responses in the aspectual and model comprehension tasks or from an examination of mean preference scores with regard to the aspectual cues (before vs. after retraining) and questions.

General Discussion

We began our experimental studies with the hypothesis that a universal linking rule of Object Affectedness is used by children to predict the syntactic privileges of verbs, but that children must learn what counts as affected. Furthermore, we argued that the same universal, in conjunction with misinterpretations of particular lexical items, gives rise to syntactic mistakes such as *I filled the water into the glass*, where the content (water in this case) is taken to be affected in the meaning of the verb. Ultimately, according to this account, the syntactic error is unlearned as the child revises his or her interpretation of the verb's meaning.

The evidence presented in this dissertation provides critical support for the non-obvious correlation of verb syntax and semantics. Our study of six common locative verbs, in Experiment 1, showed that children have a preference for content locatives (replicating Bowerman, 1982) and a bias towards the manner interpretation of locative verb meaning (extending Gentner, 1978, 1982), and that both sorts of deviations from adult language influence their learning of *fill* and *empty*. By contrast, *pour* and *dump* were rarely the sources of syntactic or semantic errors. Furthermore, this pattern of results was replicated (whenever tested) in succeeding experiments: in Experiment 2, we replicated the finding that children overgenerate the *fill*-content form and are biased (between 3;5 and 6;6) towards the manner interpretation of *fill*; in the pretesting of Experiments 3 and 4, we replicated the finding that children overgenerate the *fill*-content form, but not the *pour*-container form. On the basis of these results, we hypothesized that a general manner bias accounts not only for why children make semantic errors with *fill* and *empty* (versus *pour* and *dump*), but also--in conjunction with the linking rule--for why children prefer to overgenerate content locatives. (In this way, we unify the phenomena reported by Bowerman and Gentner.)

The specific tests of association also provided some support for the linking hypothesis. In Experiment 1, we found evidence for an association between the syntax and semantics of *empty* (esp., children biased towards the dumping manner of *empty* tended to produce more content locatives than container locatives), and for a weak association between the syntax and semantics of *fill* (esp., children between 2;6 and 5;11 who were biased towards the pouring manner of *fill* tended to produce at least one content locative). In these tests of association, we argued that the criterion of manner bias was too strong, potentially overlooking those children who had incorporated a particular manner into the meaning of a verb and yet were not biased towards that manner (e.g., children who thought that *fill* essentially means filling, but by means of pouring). In Experiment 2 we accounted for this possibility by using sensitivity as well as bias tests on the comprehension of *fill*. To our credit, we found stronger evidence for association, but again only for some of the children--those between 3;5 and 6;6. In this case, we argued that the limitation of association to this age range was consistent with the increased exposure of the older children to the *fill*-container form and with the combined sensitivity and bias results, which suggest that the older children may have regarded the pouring manner as typical of, but not essential to, the action of filling.

In general, we must stress that we tested only one plausible non-standard interpretation of filling (and emptying). It is likely that those children who attribute a manner to the meaning of *fill* show a certain amount of variation as to *which* manner they deem essential, depending upon the contexts in which they hear locatives of the verb. Indeed, the finding of *any* association is remarkable in this light.

In Experiments 1 and 2, we attempted to exploit a naturally occurring manner bias in order to show a correlation of syntactic and semantic errors. In Experiments 3-6, by contrast, we manipulated the semantics of novel verbs as an independent variable in order to show more

directly the causal nature of the linking between verb syntax and semantics. In each of these latter experiments, we were able to predict the syntactic privileges that subjects would assign to the novel verbs: we found that relatively more content locatives were produced for the verb in the manner condition (i.e., in the zig-zagging, zig-zagging or hopping, partitive, and *put*-model conditions in Experiments 3-6, respectively) and relatively more container locatives were produced for the verb in the endstate condition (i.e., in the stuffing/sagging, coloring, holistic, and *cover*-model conditions in Experiments 3-6, respectively). The results of these experiments strongly support the hypothesis that children and adults can make use of Object Affectedness in order to predict the syntactic privileges of new verbs on the basis of their meanings.

In combination, the results of Experiments 1-6 indicate that *some* sort of causal relation holds between verb meaning (cause) and verb syntax (effect). (On the possibility that syntax licenses semantics, rather than vice versa, we must conclude that this possibility seems remote in Experiments 1 and 2, where adults have presumably uttered no *fill*-content forms and where the only *fill*-content forms that a child has access to are self-generated. On the other hand, the results of Experiments 3-6 demonstrate that semantics-to-syntax linking *must* be used under some circumstances.) However, we have gathered little direct information on the scope or origin of this causal relation. Thus, our experimental evidence does not bear directly on the question of whether the correspondence is language-specific or language-general, or on the question of whether the correspondence is substantially learned or innate in origin. Nevertheless, we believe that an examination of the available evidence will favor the postulation of a universal--and by inference, innate--linking rule, which (along with other such rules) structures the correspondence between verb syntax and semantics across languages and provides a basis for lexicosyntactic productivity. In the remainder of this section, we will first defend a statement of the universal that we presented in the introduction, and then outline a

proposal of how a child may come to use the linking rule to predict the syntactic privileges of verbs.

The Universal of Object Affectedness

Our main strategy in defending this linking rule will simply be to show that the affectedness of direct objects is a universal tendency, applying across domains in English and in other languages. In other words, to the extent that this linking regularity is language-general in scope, it implies the operation of a universal linking rule. Specifically, after defining our linking rule in terms of lexical syntactic and semantic representations, we shall muster two sorts of cross-linguistic evidence: first, we will show that agent-patient verbs are universally transitive; second, we will show that linking accounts for alternations in disparate languages.

(4) an argument is encodable as the direct object of a verb if the entity to which it corresponds is affected in the meaning of the verb

In our statement of the universal (4), we assume distinct lexical representations of verb syntax and verb meaning, between which linking will occur (Rappaport and Levin, 1986; Jackendoff, 1983, 1987). The lexical syntactic representation, or *predicate-argument structure* (PAS), is assumed to consist of some indication of the number and type (e.g., Subject, Direct Object, Oblique Object) of arguments that a predicate takes in syntax. We also assume that the surface subject of unaccusative intransitives corresponds to an underlying direct object (Perlmutter, 1978; Burzio, 1986). We are agnostic as to the details of the syntactic representation, whether in terms of (e.g.) Government-Binding Theory or Lexical-Functional Grammar; these theories are intertranslateable for our purposes (see Levin 1985; Jackendoff, 1987). In order that we may be explicit about the mechanics of linking, we will follow the

convention of using variables as placeholders for the arguments in PASs (Rappaport and Levin, 1986):

(5)

a. Pour: x <y, P_{loc} z> (content-locative)

b. Fill: x <y, P_{with} z> (container-locative)

c. Load: x <y, P_{loc} z> (content-locative)

d. Load: x <y, P_{with} z> (container-locative)

In this notation, the subject (external argument) corresponds to the position filled by *x*, the direct object (internal direct argument) corresponds to the position filled by *y*, and the oblique object (internal indirect argument) corresponds to the position filled by *z*. From language to language, of course, the actual syntactic devices (word order, case and agreement marking, stress) for distinguishing between grammatical functions will vary. In English, for example, word order and case (prepositional) marking are used to distinguish direct objects (y) from oblique objects (z). What's important, however, is that *some* universal account of grammatical function be possible. A thorough discussion of representational assumptions may be found in Pinker (1989).

The lexical semantic representation is assumed to be a partial decomposition of verb meaning--a representation of the semantic elements that can be conflated in a verb's definition. According to the work of Talmy (1983) and Jackendoff (1983), among others, verb meanings across languages are organized around the concepts of motion or location in space. More abstract verb meanings, involving such domains as (e.g.) possession, emotion, and assessment are similarly organized around the motion or location of entities in an *analogue* of space--a "semantic field." Components of verb meaning accordingly specify the nature of states or events in a semantic field, including (among other things) the path, location, or orientation

of an entity; the manner in which an entity changes location or state; the causation of an event; and the means by which an event is caused. We will not, in this paper, present full representations of verb meaning; the interested reader is referred to Pinker (1989) and Jackendoff (1983, 1987). For our purposes, we can identify the argument corresponding to the content which changes location (6a) or the container which changes state (6b) in the meaning of a locative verb as a variable (y) in a substructure of semantic representation (following Rappaport and Levin, 1986):

(6)

a. ... x causes y to go into/onto z ...

b. ... x causes y to change state by means of [x causes z to go into/onto y] ...

The crucial question, of course, is what we mean for an entity to be *affected in the meaning of the verb*, as stated in (4). Until now, we have been able to use this phrase somewhat equivocally to simplify the exposition of our experimental work. At this point, we must be more precise. *Affectedness*, as we shall now use the term, merely refers to a change in the location or state of an entity (at least with respect to locatives; see our remarks on themehood/patienthood below). The linking rule, as a universal tendency, is stated in terms of affectedness. On the other hand, throughout this paper we have been using the terms *manner* and *endstate* to signify *more* than just a change of location or state in the meaning of a verb: *pour* specifies the particular manner in which content changes location; *fill* specifies the particular way in which a container changes state. If the changes of location or state specified in (6) are understood in this particular respect, linking rules should allow the language learner to "co-index" variables in the semantic (6) and syntactic (5) representations. As will become clear, determining the particular manner or endstate of a verb will involve more than just an

identification of possible affected entities. (See our remarks on the property-predicting nature of linking rules, below.)

In the terms of traditional thematic roles (which we introduce only for the sake of comparison), an affected entity is actually ambiguous between a *theme* (an entity asserted to occupy or change a location or, in the broadest construal, a state) and a *patient* (an entity 'acted-upon by an agent', interpreted very generally to mean that the patient has a role in defining what it is that makes the action of the agent an example of the verb, and not an example of a closely related verb). We will not attempt to resolve this ambiguity here, and note that most of the cross-linguistic evidence cited below satisfies both thematic roles. We would like to stress, however, that we do *not* view semantic roles as predicate-independent semantic cases (e.g. Fillmore, 1968), or linking rules as ordered lists or hierarchies of such cases, against which a form may be viewed as "canonical" or "noncanonical." Thus, we view the container and content in cases of (e.g.) loading as equally acceptable themes or patients; neither locative form is a priori canonical with respect to the other. (By contrast, learning that containers are affected in the meaning of a *particular* verb may be relatively difficult on perceptual or cognitive grounds, depending on the vagaries of input; hence our errors with *fill* and *empty*.) We note that any arguments against a predicate-independent version of innate linking do not apply to the present account (e.g., Bowerman's (in press) argument that there is no selective advantage in onset for canonical verbs).

Finally, the linking rule itself asserts that an argument is encod*able* as the direct object of a verb if the entity to which it corresponds is taken to be affected in the meaning of a verb. Our rationale for stating the rule this way--in terms of what is encodable rather than what is actually encoded--is that the affectedness of an entity does not appear to be strictly sufficient for encoding the corresponding argument as the direct object. Instead, the linking rule appears to

be property-predicting rather than existence-predicting (on this distinction, see Aronoff, 1976; Pinker, 1989), in the sense that it predicts the form that an argument *would* take, *all else being equal. All else*, in this case, refers to two sources of mediation: the operation of a *set* of linking rules (including Object Affectedness) within a domain and the clustering of the verbs of a domain into subclasses (e.g., according to particular manners or endstates). In the next section we shall argue that these mediating factors stand inbetween affectedness and direct objecthood. For the present, we shall present cross-linguistic evidence for cases in which we can successfully "hold equal" these sources of mediation; in these cases, the sufficiency of affectedness in predicting objecthood comes through (as a property-predicting regularity).

One way of seeing the influence of linking is to restrict our view to verbs where the application of the linking rules is clearest. In her review of the literature in lexical semantics, Beth Levin (1985) pays particular attention to the syntactic expression, in English and other languages, of agent-patient verbs (i.e., those in which "some generally animate entity brings about a direct (usually physical) effect on another entity" (p. 10); this construal of patient encompasses themehood, as defined above). What she finds is that agent-patient verbs are "invariably transitive in all languages" (p. 11), with the agent argument encoded as the subject and the patient argument encoded as the direct object. Examples from several domains in English include the following:

(7)

a. (causative) change of position in some manner

Sue slid the box across the floor.

Bob rolled the ball under the table.

Ted bounced the ball out of the yard.

b. change of position (in some direction)

Ken brought the wine to the party.

Sarah took the supplies to the office.

Alex carried the groceries to the car.

c. (causative) change of physical state

Toby melted the butter in the pan.

Marion hardened the candy in the pot.

Vince softened the clay in his hand.

d. physical effect

Lisa broke the vase with a hammer.

Gus crushed the can with his big tow.

Arnold smashed the window with his bat.

e. ingestion

Tim ate two pounds of chocolate.

Carol drank some gingerale.

Lou gobbled a pizza.

Counterparts to these verbs can be found in other languages, including those which are genetically, areally, and typologically distinct from English. For example, in the Australian language of Warlpiri the agent of agent-patient verbs is marked with the ergative case (ERG), corresponding to the subject, and the patient is marked with the absolutive case (ABS), corresponding to the direct object (examples from Levin, 1985; see Hale and Laughren (1983) for extended examples):

(8)

a. change of position

yirra-rni	'ERG put ABS, ERG position ABS'
ka-nyi	'ERG carry ABS, ERG transport ABS'
rarra-ka-nyi	'ERG drag ABS'

b. physical effect

yurrpa-rni	'ERG grind ABS' (as seed, ochre)'
yarlki-rni	'ERG bite ABS'

Notice that the precise interpretation of affectedness varies from one domain to the next, sometimes involving a manner or direction of motion and sometimes involving a change of state. In the case of agent-patient verbs, the membership of a verb in one domain or another may have relatively little effect on its transitivity; according to Levin, the syntactic expression of agent-patient verbs (holding constant the number of arguments) is uniform within and across languages. For two-argument verbs which deviate from the agent-patient standard, however, Levin argues that there is greater variability in syntactic expression, both within and across languages. In these cases, different verbs within a given domain may have different syntactic properties (9a,b) or the same verb may have alternative syntactic expressions involving a change in transitivity (9c):

(9)

a. emotion: experiencer as subject

Adam loves Eve.

Cain hates Abel.

Pam fears Hurricanes.

emotion: stimulus/agent as subject

Hurricanes scare Pam.

Lectures bore Fred.

Ornithology thrills Herbert.

b. perception: transitive

Boris saw the sign.

Laurie heard the siren.

perception: intransitive

Boris looked at the sign.

Laurie listen to the siren.

c. surface contact: transitive

Ken hit the wall.

Lois slapped Clark.

Garry struck the chair.

surface contact: intransitive

Ken hit at the wall.

Lois slapped at Clark.

Garry struck at the chair.

Similar examples, also from relatively abstract semantic domains, include verbs of cognition (*think, doubt, occur to*), desire (*want, prefer, hanker after*), and assessment (*esteem, value, prize*) (See Levin, 1985; Talmy, 1985). Across languages, according to Levin, these same classes show syntactic variability. She gives the example from Warlpiri of *paka-rni* 'strike', for which the contacted entity may take the absolutive or dative case (cf. the change in transitivity in (9c)). In some languages, in fact, verbs of surface contact may "display variation as to whether they are among the transitive verbs in the first place, and if so, whether they allow more than one way of expressing their arguments" (p. 12). By contrast, English speakers cannot say **John slid at the box* (meaning that John intended to slide the box; cf. 9c), and similarly for the agent-patient verbs of other languages. Given the pattern of little variation in the syntactic expression of agent-patient verbs, versus more variation in the syntactic expression of verbs falling outside of this class, Levin concludes that if a verb belongs to the agent-patient class, it must have a transitive expression (with the patient encoded as the direct object), but not necessarily the converse.

A stronger conclusion consistent with this pattern of results, though still short of necessity, is that the uniformity in the syntactic expression of agent-patient verbs reflects a true universal *tendency* in the linking of verb meaning and syntax. Affectedness appears to serve an organizing role for the expression of direct objects; at least, no *other* sufficient condition on the meanings of verbs has been discovered to have such an obvious codifying force on the expression of direct objects. Support for this stronger conclusion comes from the thorough cross-linguistic study of transitivity by Hopper and Thompson (1980). On the basis of their findings, Hopper and Thompson claim that transitivity--universally--involves a number of correlated morphosyntactic and semantic components, including affectedness. In a nutshell: high transitivity is associated with telic, punctual events in which a potent agent acts volitionally upon an individuated, affected patient. In particular, they conclude that "partitive

O's [partially affected entities] are universally associated with intransitive verbs, or at least with some signal of *reduced* transitivity" (p. 263). The cross-linguistic evidence reviewed by Hopper & Thompson (1980) and Levin (1985) suggests that linking doesn't exist (solely) for the purpose of allowing children to predict the syntactic privileges of verbs, but serves as a conservative influence on the syntactic expression of arguments--perhaps as a means of insuring that the general information of *who* did *what* to *whom* is not lost over generations of language change. (The potentially far-reaching consequences of language without linking rules are difficult to fathom: paradoxically, without the stabilizing and yet productive force of linking rules, lexical change might be too fast for its own good, or not possible at all.)

In a final set of examples, the influence of linking will be made apparent not because the application of the linking rules is univocal, but because it admits of variation (sometimes involving a change in transitivity, and sometimes not). A perfect example of this type is the locative alternation. Superficially, it might be argued that the existence of alternations such as the locative argues *against* the linking of semantic and grammatical relations--after all, in one form a content is a direct object, and in the other form a container is a direct object. However, a deeper level of analysis, supported by our experimental research, reveals that the direct object in these cases corresponds to an affected entity in the meaning of the verb. As we have already seen in (7) and (8), this correspondence is apparent across other constructions (which also differ superficially, according to the domain of the verb).

Most relevant here is that the locative forms of a given verb differ in their implications, a phenomenon that has been labelled the holistic interpretation (Anderson, S., 1971; Schwartz-Norman, 1976); for example, *John loaded the cart with the apples* implies that the capacity of the cart has been exhausted, but *John loaded the apples into the cart* does not. In this context, we can attribute the holistic interpretation to the application of the linking rule(s) to different

arguments. More generally, we will show that in cases where a particular predicate and its arguments admit of more than one syntactic expression, and in which an argument is alternatively encoded as the direct or oblique object, a subtle semantic difference usually, if not always, accompanies the alternation. Because the only difference between the forms is in the linking of arguments to grammatical relations, and not in the nature of the arguments themselves, any semantic differences must be directly relevant to the issue of linking.

One methodological note before we proceed. In the following examples of cross-linguistic evidence, we have limited ourselves to alternations between direct and prepositional objects in which the alternative forms differ in their implications. We make no pretensions, of course, that the sampling is anywhere close to exhaustive (especially since thorough cross-linguistic evidence is not available), but we do feel that it is representative, given two provisos. First, we have ruled out alternations which involve the addition/deletion of any elements other than the relevant case or agreement markers (e.g. negativity markers), on the grounds that the meanings of these additional elements might be sufficient, but not necessary, to force the alternation. Second, we have ruled out alternations (esp. of morphological case) which involve semantic properties of the direct/oblique object itself (versus semantic properties of the verb). Such properties often enhance the individuation of the object, where *individuation* refers to the distinctness of an affected entity (especially, from the agent), and encompasses such properties as animacy, humanness, definiteness, numericity, and count/mass status, among others (Hopper and Thompson, 1980). We acknowledge that, in some cases, morphological case marking may do other things besides distinguishing grammatical function (Comrie, 1981). We suggest, however, that individuation and affectedness are not independent properties, but that the affectedness of an entity may presuppose a high degree of individuation. (Individuation, as well as perspective, has a subtle effect on locatives. For example, we find the following sentence acceptable: *John was loading carts with apples for hours, but never managed to fill*

any of 'em. But compare: ?*John loaded the cart with apples for hours, but never managed to fill it.* (We owe this example to Robert Van Valin.))

Consider the pairs of sentences in (10): (a) the conative; (b) the "locomotive"; (c) the dative; and (d) the locative. Each of these pairs illustrates how an argument of a verb may be alternatively expressed as a direct object or as a prepositional object, with a subtle semantic "shift" accompanying the alternation. Specifically, the argument in question (italicized in 10) may correspond in the meaning of the verb to either an affected entity, if linked to the direct object, or to a reference object (most commonly, a goal, source, or location), if linked to the oblique object. Notice that the precise interpretation of affectedness depends upon the domain in question. In (a), the first sentence of each pair (with the NP as direct object) implies the successful contact (*slap*, or *hit* (in 9c)) or penetration (*cut, slash*) of the affected entity. By contrast, the second sentence of each pair (with the NP as the object of *at*) only implies the *intent* to act upon a goal (cf. *Kurt cut the bread into two pieces*; ?*Kurt cut at the bread into two pieces*). Similarly, in the case of verbs of locomotion (10b), the entities may be interpreted as "conquered" (e.g., ?*Larry leapt the chasm, but fell short of the other side*) or as reference objects (cf. *Larry leapt over the chasm, but fell short of the other side*). In the case of three-argument forms, the same semantic shift obtains. In the dative examples with *teach* (10c), the double-object (first), but not the prepositional (second) form implies that the children have actually learned Spanish (Green, 1974). The different implications of these two sentences seem especially clear, perhaps because the activity of teaching (which does not necessitate any learning) is so often dissociable from its accomplishment. In fact, as was explicit in the design of Experiments 2, 5, and 6, the affectedness of a reference object in the meaning of a verb is conflated with the ability of the verb to take an accomplishment (or achievement) reading--that is, the ability to specify a definite endpoint to the action. Accordingly, all of the verbs in (10) may be alternatively viewed as specifying either accomplishments/achievements or activities, in

the first and second sentences, respectively. (See Tenny (1988) on the relation of affectedness and verb aspect.)

(10)

a. Kurt cut *the bread*/at *the bread*
   Bill slashed *the tire*/at *the tire*
   Rachel slapped *the stranger*/at *the stranger*

b. Larry leapt *the chasm*/across *the chasm*
   Betty swam *the channel*/across *the channel*
   Ted climbed *the mountain*/up *the mountain*

c. Jake taught *the children* Spanish/Spanish to *the children*
   Joe threw *Frank* the ball/Joe threw the ball to *Frank*
   Pam told *Sue* the secret/told the secret to *Sue*

d. John sprayed *the wall* with paint/paint on *the wall*
   Cathy loaded *the cart* with apples/apples onto *the cart*
   Max smeared *the mirror* with paste/paste onto *the mirror*

A final comment on these examples from English is that the semantic shift is demonstrably a product of *two* changes in the linking of semantic and grammatical functions: an argument is no longer linked to one grammatical function (e.g., oblique object), and is now linked to another (e.g., direct object). (Note that the case of three-argument forms is complicated by the simultaneous switch in linking for the other non-agentive argument. It is unclear what semantic effects follow from the recoding of the other argument.) Although we have focused on the

linking of direct objects, consider a second linking rule: that goals, sources, and locations tend to be encoded as oblique objects (Talmy, 1983; Jackendoff, 1983). The consequences of this linking rule can be seen in (10); the spatial properties of reference objects are "collapsed" upon their recoding as direct objects. This is quite clear in the case of the locatives (10d), and may give some insight into why the affectedness of containers is typically holistic (versus some other possible interpretation of affectedness, such as 'coming into contact with some content'), as if the three- or two-dimensional geometry of a container or surface were being reduced to the one-dimensional geometry of an affected entity (see Talmy, 1983, on the geometry of themes and reference objects). In the next section, we shall consider the consequences of the fact that an ensemble of linking rules operate together.

In languages which are genetically and areally distinct from English, we find not only the same alternations between direct object and oblique object, but also the same "semantic shifts" which accompany them. We have already mentioned the example from Warlpiri of *paka-rni* 'strike', for which the contacted entity may take the absolutive or dative case. According to Levin (1985), the absolutive (but not dative) marking implies that a surface has been contacted. As in English, the conative alternation in Warlpiri also applies to verbs of penetration (*nominative* in this example corresponds to the direct object; Hale, 1973, cited in Moravcsik, 1978):

(11)

njuntululu npatju pantuṇu ŋatju
"you-erg. you-I spear-past I-nom."
'you speared me'

njuntululu npɑtjuḷa pantuṇu ŋatjuku

"you-erg. you-I-? spear-past I-dative"

'you speared me' or 'you tried to spear me'

In her cross-linguistic study of the case marking of objects, Moravcsik (1978) claims that a semantic shift also accompanies the locomotive, dative, and locative alternations in Hungarian. Despite the fact that Hungarian is genetically and areally distinct from English, the differences between the pairs of Hungarian forms in (12) and (13) precisely parallel the differences between the corresponding English translations. In these examples from Moravcsik, only the first sentence of each pair (and its translation) carries the implication that a reference object is affected in the meaning of the verb (here, the accusative (acc.) marks the direct object):

(12)

a. Megmaszta a hegyet

"up-climbed-he/she-it the mountain-acc."

'he/she climbed the mountain'

Felmaszott a hegyre

"up-climbed-he/she the mountain-onto"

'he/she climbed up the mountain'

b. Atugrotta az arkot

"across-leapt-he/she-it the ditch-acc."

'he/she leapt the ditch'

Atugrott az arkon

"across-leapt-he/she the ditch-on"

'he/she leapt across the ditch'

c. Atuszta a tavat

"across-swam-he/she-it the lake-acc."

'he/she swam the lake'

Atuszott a tavon

"across-swam-he/she the lake-on"

'he/she swam across the lake'

(13)

a. Janos bemazolta a falat festekkel

"John in-smeared-he-it the wall-acc. paint-with"

'John smeared the wall with paint'

Janos ramazolta a festeket a falra

"John onto-smeared-he-it the paint-acc. the wall-onto"

'John smeared paint on the wall'

b. Janos beultette a kertet fakkal

"John in-planted-he-it the garden-acc. trees-with"

'John planted the garden with trees'

Janos elultette a fakat a kertben

"John away-planted-he-them the trees-acc. the garden-in"

'John planted the trees in the garden'

According to Rappaport and Levin (1986), the holistic effect also appears in every other language (besides Hungarian and English) in which the locative alternation has been studied. Examples similar to (13) may be found in Berber (Guerssel, 1986), French (Boons, 1971), Japanese (Fukui et al., 1985), Kannada (Bhat, 1977), and Russian (Veyrenc, 1976) (citations from Rappaport and Levin, 1986).

Besides the fact that the verbs in (12) and (13) are marked in transitive sentences for agreement with accusative NPs, notice that they are also prefixed (i.e., with *meg-* and *at-* in 12 and *be-*, *ra-*, and *el-* in 13). According to Moravcsik, these verbal prefixes are similar to the post-verbal particles in English (e.g., *up*, *down*, *through*) in that they can indicate either directionality or completedness, or both. It turns out, however, that there is a striking difference between English and Hungarian locatives both in the use of these particles and in the ability of locative verbs to alternate. In the case of English locatives, the use of particles is optional and furthermore restricted to the container-locative form (Fraser, 1971); as we have seen, the completive particle *up* in container locatives serves to emphasize the change of state that a container undergoes. In the case of Hungarian locatives, particles are regularly (if not obligatorily) prefixed onto verbs in both locative forms, where they appear to serve more than just an emphatic function. In fact, *all* Hungarian locative verbs (of addition) are alternators (Moravcsik, 1978, n. 2), but the ability to alternate appears to depend crucially on the presence of the prefixes. In this regard, Rappaport and Levin (1985) make the following observation: "the counterparts of *fill* and *pour* in both Russian and Hungarian are alternating verbs, but in these languages the verbs in the two variants, while containing the same root, differ in aspectual prefixes" (p. 38, n. 20).

In support of the claim that the ability of a verb to alternate depends crucially on its meaning, we can cite other languages (besides Russian and Hungarian) in which the locative alternation is accompanied by changes in verbal morphology that explicitly indicate the affectedness of the container as a function/extension of verb meaning. Fukui, Miyagawa, & Tenny (1985) have noted that Japanese and English differ both in their degree of locative alternation and in their degree of productive verbal morphology, and that these two factors are related: although Japanese (ordinarily) allows fewer locative verbs to alternate than does English, the addition of the "holistic" morpheme *-tsukusu* ('exhaust') to Japanese verbs greatly increases the number of alternators. A similar phenomenon occurs in the Nigerian language of Igbo; verbs must be compounded before accepting the equivalent of the full container locative (Nwachukwa, 1987). These observations underscore the *systematicity* of the locative alternation across languages, and in particular argue that the non-alternation of *pour* and *fill* in English is not an arbitrary gap, but rather a (systematic) consequence of verb meaning.

In summary, two principal findings support the universality of Object Affectedness. First, the uniformity across languages with which agent-patient verbs express their arguments reflects a universal tendency for direct objects to correspond to affected entities. Second, the finding that semantic shifts involving affectedness accompany alternations in disparate languages suggests that affectedness is a consequence of universal direct-object linking, and that the ability of verbs to alternate is a consequence of their meanings. To the extent that the linking regularity of object affectedness is universal, we argue that it is innate.

Inversely, we question whether any story in which linking regularities are learned (e.g., Bowerman, 1982, in press) can account for the observed universal tendencies. Furthermore, as we have argued throughout, learning can't plausibly account for fine-grained semantic errors with *fill* and *empty*, or resolve Baker's Paradox, without attributing to children a fairly

complex (unlearnable) mechanism for linking. Finally, we have demonstrated in Experiments 3-6 that affectedness, not the (learnable?) predominance of content locatives per se, correctly predicts the syntactic privileges that children grant novel verbs differing in manner or endstate, or differing in the partitive/holistic extent of effect upon a container.

Using the Linking Rule

Turning to the question of *how* the linking rule is used by children, it is undoubtedly *not* the case that children have direct access to a linking rule, which they use in a mechanical fashion to check off the syntactic privileges of verbs. In particular, there are two sources of "mediation" which appear to prevent the direct application of the linking rule to the lexical semantic representation of a verb: the operation of a set of linking rules (including Object Affectedness) in a domain and the clustering of the verbs of a domain into subclasses. Together, these two complicating factors suggest that the sufficiency of affectedness in predicting objecthood, as demonstrated above, is property-predicting in nature, not existence-predicting. In this subsection, we shall discuss each source of mediation in turn.

We assume that all of the arguments of a verb must be assigned to grammatical functions, and that each argument must be assigned to a unique grammatical function in a sentence, and conversely (i.e., the "Theta-Criterion" of Chomsky, 1981; the "Coherence" and "Completeness" Principles of Bresnan, 1982). (We note that although there are languages like Japanese where virtually any argument of a verb may be unexpressed, the information must be supplied by context (Fukui et al, 1985).) Therefore, we must consider the possibility that a linking rule such as (4) does not operate in isolation, but rather works in concert with linking rules for the subject and the oblique object. In the case of the subject, which has been the focus of most of the cross-linguistic work on grammatical relations, the available evidence supports the existence of linking: there is a universal tendency for the agents and causal forces of actions

to be encoded as subjects (Keenan, 1976). Less systematic work has been done relevant to the linking of oblique objects, but what work has been done suggests that most of the entities encoded by oblique objects can assimilated to reference objects of some kind (Talmy, 1983; Jackendoff, 1983). Furthermore, Pinker (1989) points out that a linking rule for oblique objects need not co-opt all of the differences between prepositions or oblique case markers, since the semantic and syntactic properties of these markers are surely specified in their own lexical entries. On the basis of the available evidence, therefore, we shall assume that several linking rules operate together. Rappaport & Levin (1986) present the following as a typical set of linking rules (based on Carter, 1976a,b; Ostler, 1980; see Pinker, 1989, for a more precise formulation of linking rules):

(14)

a. Link the agent role with the external argument [subject] variable in the predicate-argument structure.
b. Link the theme or the patient role with the direct argument [direct object] variable in the predicate-argument structure.
c. Link each remaining theta-role to an indirect argument [oblique object] variable in the predicate-argument structure which is associated with an appropriate preposition.

Our intention here is not to justify these particular statements of the linking regularities, but rather to ask whether predicate-argument structure is another level of lexicosyntactic representation, besides (subsumed) grammatical function, at which we must capture generalizations about linking. Notice that this is not the question of whether a set of linking rules must operate together (which we assume above); obviously, it would be somewhat meaningless to assume that the linking rules operate in isolation given that the arguments of a verb must be expressed together and that other linking rules plausibly exist. Instead, we want

to know whether children, in learning the syntactic privileges of a verb, perform linking *to established PASs* rather than to a set of grammatical functions. We shall call the ability/necessity of children to use linking rules in this fashion *PAS-linking*. Our question, then, is this: is there any evidence for PAS-linking? The answer to this question in undoubtedly *yes*; in particular, the available evidence suggests that children have and use foreknowledge about the *possible* range of PASs of a new verb on the basis of its membership in a semantic domain. Roughly speaking, the story here is that a child learns the syntactic privileges of (e.g.) a locative verb (of addition) by first attending to the range of forms which are commonly used in a language to express events about adding content to a container, and by then choosing--within this constrained space of possibilities--one locative form or the other, or both, on the basis of (PAS-) linking.

In the introduction, we characterized a *domain* as a set of verbs with shared semantic and syntactic properties. We have already discussed examples from several domains, including locative verbs of addition, locative verbs of removal, (dative) verbs of giving, verbs of emotion, verbs of perception, verbs of motion, verbs of change of state, verbs of physical effect, and so on. Although some readers may regard the concept of a *domain* as hopelessly fuzzy, we maintain that some such notion is functional in language acquisition, and in fact widespread in the literature of lexical semantics and language acquisition. (It remains for future cross- and psycholinguistic study to ferret out the precise boundaries of these domains, as well as the places where they overlap.) Assuming a workable notion of domain, we can show that the domain of a verb can be used, in principle, to predict the possible range of the verb's syntactic privileges. The strongest evidence in support of this claim comes from the observation that the verbs of a given domain, when they can be identified across languages, tend to express their arguments in the same limited number of ways, and differently from verbs in other domains. As Levin (1985) puts it, "Even in [semantically coherent] classes that allow

alternate realizations of arguments, the patterns attested seem to be limited. These possibilities generally appear to be drawn from a set of options available to verbs in a given class across languages" (p. 13). Although thorough evidence on the realization of particular semantic classes across languages is not available, in those languages where the locative alternation has been studied, the predicted result obtains: "when a language manifests the [locative] alternation the verbs that participate in the alternation fall into the same broad semantic class as the English locative alternation verbs" (p. 36, Rappaport & Levin, 1985). Earlier, in fact, we cited evidence that the locative alternation in a number of languages involves not only the same semantic class of verbs, but the holistic effect as well. Finally, some cross-linguistic evidence supports the view that the sets of PASs for different domains are underlyingly different despite the linking rules that they share; for example, Russian displays the locative, but not the dative alternation, whereas English displays both (Rappaport and Levin, 1985). This sort of variation across languages also suggests that the total set of PASs may vary from language to language.

On the distinctness with which verbs of a domain are expressed, Levin (1985) argues that the child must be sensitive to certain semantically coherent classes of verbs. Within English, we have already cited evidence on differences in the syntactic expression(s) for a number of classes, including: verbs of ingestion, physical effect, change of position, or change of state; versus verbs of emotion; versus verbs of surface contact; versus verbs of perception; versus verbs of locomotion; versus locative verbs of adding; versus dative verbs. Furthermore, Levin presents cross-linguistic evidence that we can further subdivide the class of agent-patient verbs; for example, across English, Warlpiri, and French, verbs of change of position (e.g., *slide*) and change of physical state (e.g., *melt*), as well as change of psychological state (e.g., *amuse*) are separated in the devices that they use to mark the "anti-causative" alternation: in English, the former two classes pattern together (i.e., *the box slid*; *the butter melted*; *the children *amused/were amused*); in Warlpiri, the latter two classes pattern together; and in French, all

three classes pattern together. Many other examples could be given. In summary, if two verbs belong to different domains, it is a good bet that they have different ranges of syntactic privileges.

It might be argued (though not without some difficulty) that the similarity in the expression of verbs of a particular domain across languages, and the distinctness of their expression from verbs of other domains, is simply due to the operation of independent linking rules working in a constrained cognitive space. However, it is more difficult to account in a similar fashion for several psycholinguistic observations. First, we noted earlier that children rarely use the verb of one domain to assert the occurrence of an event usually expressed by the verbs, and in the PAS(s), of another domain. For example, children rarely, if ever, use *eat* in a locative form (e.g., ***I ate a spoon into my pudding***), although we might expect such errors on the basis of individual linking rules (or on the basis of unconstrained syntactic rules, as discussed by Bowerman, 1982). That children don't make mistakes across domain boundaries, even when similar semantic roles are involved, shows that they are, in practice, sensitive to the PASs of a domain. Second, Levin (1985) notes that lexical *extensions* occur in adult speech, in which the verbs of one domain (e.g., the manner of motion verb *slide*; the surface-attachment verb *sew*) may be used to express the *means* of an event which is itself usually expressed by the verbs, and in the PAS(s), of another domain (cf. change of possession; creation). Thus, we may say ***John slid Mary the package*** and ***John sewed the remnants into a shirt***, respectively. Far from arguing against domain-specific PAS-linking, we take the intuition of lexical extension to be an important datum in favor of our claim: speakers who use such extensions must rely on the ability of the hearer to associate a form with a particular domain. In the example of the lexical extension of *slide* in the double-object dative, a component of meaning (change of possession) is neither conflated in a verb's definition (nothing about the meaning of *slide* specifies transfers of possession) nor supplied by individual linking rules (none of which are domain-specific),

but rather emerges from a combination of linking rules devoted to a particular domain of verbs. In particular, we have elsewhere demonstrated (for English-speaking children and adults) that the acquisition and use of double-object datives is generally constrained to express the change of state that a possessor undergoes (Gropen et al, 1989).

On the basis of the cross-linguistic and psycholinguistic evidence, therefore, it seems likely that two (additional) steps are relevant in learning the syntactic privileges of verbs, besides just applying the linking rule of Object Affectedness to the meaning of a verb: first, children appear to narrow down the range of possible syntactic expressions of a verb on the basis of its membership in a domain. That is, children have to learn which PASs go with which domains, even though the matching may be constrained by individual linking rules to a great extent. Second, once the PASs of a domain are known, the joint satisfaction of linking rules may provide a *necessary*, but not *sufficient* condition on the ability of a verb to take the corresponding PAS. In other words, our claim here is that linking regularities may be stated more strongly at the level of PASs within a domain than at the level of individual grammatical functions. Notice furthermore that this condition is still formulated as a property-predicting rule (since not every verb in a domain may be expressed in every relevant form), but presumably a more powerful one. (As an aside, we interpret the "syntactic bootstrapping" hypothesis of Landau and Gleitman (1985) as follows: if the set of PASs, but not domain, of a particular verb is known, and furthermore if that set of PASs is known to correspond to a particular domain, then the verb may be inferred to belong to that domain.)

The hypothesis of PAS-linking is a claim about how the domain of a verb constrains the lexicosyntactic side of the linking equation. We can also address the question of what is the appropriate level of lexico*semantic* representation at which to capture generalizations about linking: is it the case that the PASs within a domain are associated with stable semantic

representations which are relevant to linking? This is a difficult question to answer, primarily because we have no direct evidence on the issue. Notice that the finding that semantic shifts accompany alternations, as discussed in the last section, isn't really evidence for the existence of a stable semantic correlates of PASs; we argued there that the holistic effect could be traced to the re-linking of one argument--the "promotion" of the container argument from oblique object (i.e., goal) to direct object (i.e., affected entity). Furthermore, although it may be true that the *choice* of PASs for a verb is domain-specific, it isn't necessarily the case that the *interpretation* of putative semantic correlates of PASs is domain-specific. Thus, the domain of a V cannot be consistently inferred solely from its acceptance of PASs such as (e.g.) NP-V-NP-*into*-NP, NP-V-NP-*with*-NP, and NP-V-NP. For example, NP-V-NP can be about a change of position (*John slid the box*), a change of physical state (*John melted the butter*), a change of psychological state (*John scared the boy*), an act of perception (*John saw the boy*), and so forth. We may certainly be able to narrow down the likely interpretation of *some* PASs (especially the double-object form), but the critical factor seems to be the range of uses of a PAS in a particular language, and not the existence of a stable semantic correlate of that PAS.

On the other hand, the postulation of stable semantic representations corresponding to particular PASs is not without its theoretical merits (as we shall discuss below), and by making such an assumption we aren't begging any crucial questions about how linking rules are used. We shall therefore follow the work of Pinker (1989) in using the term *thematic core* to designate the composite semantic representation corresponding to a particular PAS. Thematic cores are lexical semantic representations which are independent of particular predicates, but--as we present them here--tied to a particular domain of arguments. (Pinker (1989) hypothesizes that thematic cores are not inherently domain-specific, but function across domain boundaries. We will not, in this paper, pursue this hypothesis.) For locatives involving the addition of content to a container, we shall assume the following thematic cores (the general structure of

these thematic cores is similar to, but abstracted from, the particular definitions of locative verbs given in (6), from Rappaport and Levin (1986)):

(15)

a. the thematic core 'X causes content Y to go into/onto container Z' corresponds to the content-locative PAS 'x <y, P_{loc} z>'

b. the thematic core 'X causes container Y to change state by means of [X causes content Z to go into/onto container Y]' corresponds to the container-locative PAS 'x <y, P_{with} z>'

In this view, the child never applies an individual linking rule such as Object Affectedness directly to the meaning of a verb, but instead uses the thematic cores of a domain to determine the possible ways in which the set of linking rules may be satisfied for that domain; the matching of a verb to a thematic core is a necessary condition for the verb to take the corresponding PAS. In practice, the pair of thematic cores corresponding to alternative PASs in a domain may serve to structure the domain into two broad divisions (e.g., manner (15a) and endstate (15b)), and thus may guide the child's decisions about how to express the verbs of a domain. We shall return to the critical question of precisely how a child determines the sufficiency of a verb to accept a particular PAS within a domain, but let us briefly consider what the notion of a thematic core buys us. First, the linking of locative forms to semantic representations gives us a unified account of the (near) paraphrase relation between locative forms, and of the ability of some verbs to alternate between them. The thematic cores in (15) satisfy the paraphrase requirement (see Rappaport and Levin, 1986) by virtue of the

substructure that they have in common, 'X causes content Z(Y) to go into/onto container Y(Z)'. In fact, this paraphrase relation is more parsimoniously captured at the level of the thematic core than at the level of individual verb meanings (as in Rappaport and Levin, 1986). Furthermore, because the *set* of thematic cores of a domain specifies the range of syntactically relevant ways in which the verbs of that domain may be interpreted, any pair of thematic cores constitutes a "broad-range" rule that can be used to predict the possible alternation of verbs in that domain. Thus, the notion of a lexical rule, though one that is semantically based, falls out of this account (Pinker, 1989).

Second, the postulation of domain-specific thematic cores allows us to systematize the knowledge of a domain or *argument space* that language learners appear to have. In Experiment 2, we showed that sensitivity to manner or endstate (qua achievement) doesn't necessarily accompany children's early use of the verb *fill*. An interpretation consistent with all of the available evidence is that before a child fixes the meaning of a verb in terms of manner or endstate, he or she must have learned that contents and containers are the appropriate arguments over which to define the meaning of a verb. (Encoding domain in a thematic core also explains the phenomenon of lexical extension without having to resort to special rules of interpretation.) This requirement is captured in our explicit reference to contents and containers in (15). Here, *content* and *container* should be viewed as labels, ultimately cashed out in terms of semantic primitives which specify the topological properties of potential containers and contents (Talmy, 1983; Jackendoff, 1983). These properties, among others, will be prominent in our discussion of conflation classes, below. In sum, the child must have experience with the particular domains of human activity that the verbs of a language make reference to, domains which may then be reified in terms of thematic cores and their associated PASs.

Finally, we can also provide a more specific answer to the question of what linking rules are for. Earlier, we suggested that linking doesn't exist (solely) for the purpose of allowing children to predict the syntactic privileges of verbs, but serves as a conservative influence on the syntactic expression of arguments. We can now suggest the particular function that linking rules might have: linking rules, together with the PASs in a domain, can be used to triangulate on the language-specific thematic cores within an argument space. This division of labor, between universal linking rules and language-specific thematic cores, is consistent with the cross-linguistic evidence (esp., the differences between English and Russian) that we reviewed above.

We now turn to the second source of mediation between affectedness and objecthood: the clustering of the verbs of a domain into subclasses. Even if we assume that a child can accurately classify a verb according to its domain, we are still faced with a learnability problem. In fact, the logic of Baker's Paradox that we introduced in the General Introduction assumes that "local errors" of overgeneration are the real threat to learnability; for example, a child may learn that some locative verbs of addition (e.g., *load*) can be expressed in two forms, and on that basis extend the privilege to other *similar* verbs (e.g., locative verbs such as *pour* and *fill*). Without recourse to feedback about which sentences are *not* in the language, the child would then be unable to unlearn any overgenerations (e.g., **John poured the glass with water*; **John filled water into the glass*). In our present terms, a child might even predict that every locative verb of addition was an alternator on the basis of the broad-range rule in (15). How, then, does the child learn the correct syntactic privileges for *pour*, *fill*, and *load* on the basis of linking rules?

In order to explain why *pour*, *fill*, and *load* don't all have the same syntactic privileges, Pinker (1989) has hypothesized that a child determines the sufficiency of a verb to accept a

particular PAS (within a domain) on the basis of its membership in a semantically cohesive subclass. Linking rules serve as the basis for this process in the following way: the pair of thematic cores corresponding to alternative PASs in a domain structure the domain into two divisions (e.g., manner (15a) and endstate (15b)), and thus guide the child's search for relevant subclasses. Crucially, the mere themehood of a content or container argument, according to (15a,b), is not enough; the child needs independent evidence that a *particular* manner or endstate is incorporated into the meaning of a verb. Finally, if a verb (e.g., *fill*) is found not to belong to any subclass of a given type (e.g., manner), the child reaches the conclusion that the verb is incompatible with the relevant ("parent") thematic core and its associated PAS (unless the verb is heard in that PAS in the positive input).

Our goal here will be to present a list of subclasses, adapted from Pinker (1989), which could in principle be used by children in learning the syntactic privileges of locative verbs of addition. (Although we will restrict our remaining discussion to these verbs, a similar analysis has been performed for locative verbs of removal (Pinker, 1989); in addition, we note that some manner verbs (i.e., non-directional) may belong to both domains.) The method we employed was to search the fairly exhaustive list of 126 locative verbs in the Appendix (from Rappaport and Levin, 1985) for the dimensions of verb meaning which provided for the most natural semantic clustering of the verbs, while still accounting for their syntactic distribution. By *natural semantic clustering* we mean that we attempted to find dimensions that minimized the semantic distance between verbs which shared a value on that dimension (many of them differing only in dialect, register, or connotation) and that maximized the semantic distance between verbs in different clusters or subclasses. The dimensions themselves were divided into two broad groups on the basis of whether they involved the affectedness of the content (i.e., a *manner dimension*) or the affectedness of the reference object (i.e., an *endstate dimension*). This broad separation of dimensions into manner versus endstate follows from the hypothesis

that the pair of thematic cores in (15) literally guide the child's search for relevant subclasses. Our analysis provides for a potential solution to the learnability problem in the following way: if a verb belongs to at least one manner subclass, it will accept the content-locative form; if a verb belongs to at least one endstate subclass, it will accept the container-locative form; and if a verb belongs to at least one manner subclass *and* at least one endstate subclass, it will accept both locative forms. (Note that the method in Pinker, 1989, differs from that presented here. Among other differences, the approach there was to formulate subclasses for which all of the members either alternate or do not alternate.)

The analysis of locative verbs of addition into manner and endstate subclasses is presented in Table 32. The major subheadings (i.e., *Spatial Distribution of Content*, *Geometry of Reference Object*, and *Purpose*) correspond to dimensions along which included verbs may vary according to subclass or cluster. The minor subheadings correspond to the following clusters of verbs: particle/blob, array of particles/blobs, continuous stream, layer, vertical mound, compressed mass, circle/coil, suspension, edge, surface (coverage), surface (distribution of a set of objects), surface (support), layered medium, co-extensive medium, container, path, function, forceful surface contact, qualitative/esthetic (positive), qualitative/esthetic (negative). Before we discuss each of these particular subclasses, let's be clear about what is being claimed. We have proposed dimensions which are as general as possible (though not every endstate verb need specify a value for one dimension if the other dimension is relevant), and for which different clusters of verbs specify different discrete values along that dimension. Notice that this is a fairly minimalistic view of the correspondence of verb syntax and semantics (assuming, of course, that there is a correspondence) in that narrowly circumscribed aspects of meaning are taken to be relevant to the syntax of a verb. Specifically, the manner verbs (i.e., the verbs belonging to manner subclasses) *must* specify

Table 32

Manner and Endstate Subclasses of Locative Verbs of Addition

Manner Dimension

Spatial Distribution of Content (as it changes location):

particle/blob: dribble, drip, slop, slosh, spill

array of particles/globs: drizzle, scatter, shake, bestrew, spatter, sprinkle, shower, dust, splash, splatter, spray, sow

continuous stream: ladle, pour, dump, spew, inject, squirt, flood

layer: dab, daub, smear, spread, smudge, plaster, slather (resulting layer may vary in shape and evenness)

vertical mound: pile, stack, heap

compressed mass: wad, cram, crowd, jam, stuff, pack

circle/coil: coil, spin, twirl, twist, whirl, wind, wrap

suspension: hang, string, drape

Endstate Dimensions

Geometry of Reference Object (as it becomes holistically affected):

edge: edge, trim$_1$, string$_1$

surface (coverage): blanket, cover, inundate, line, shroud$_1$, vest$_1$, bandage$_1$, coat, deluge$_1$, douse, encrust, face, inlay$_1$, pad, pave, plate, tile, daub$_1$, spread, dust, spray$_1$, flood, wrap, drape$_1$, plaster, slather

surface (distribution of a set of objects): litter$_1$, spot$_1$, stud, blot$_1$, clutter$_1$, riddle$_1$, splotch, dapple, bestrew, spatter$_1$, splatter$_1$, sprinkle, splash$_1$, sow

surface (support): burden$_1$, pile, stack, heap

Note: Alternators are underlined; non-alternators are presented in plain text. Subscripts indicate the membership of a verb in two endstate subclasses (but nothing is implied by the pairing of subscripts to readings). The verbs in this analysis are from Rappaport and Levin (1985) (see the Appendix).

Table 32 (continued)

Manner and Endstate Subclasses of Locative Verbs of Addition

Endstate Dimensions (continued)

Geometry of Reference Object (as it becomes holistically affected): (continued)

layered medium: lard$_1$, interlace, interlard, interleave, intersperse, interweave, vein, ripple

co-extensive medium: drench, saturate, suffuse, imbue$_1$, impregnate, infuse$_1$

container: fill, (re)populate, cram, crowd, stuff, pack$_1$

path: block, choke, clog, dam, plug, smother$_1$, stop up, bind, entangle$_1$, lash, jam

Purpose (in terms of abstract changes of state of the reference object)

function/potential (the reference object is enabled to perform its function): stock, load, pack$_2$

forceful surface contact: dab, squirt, spray$_2$, splash$_2$

qualitative/esthetic (positive): adorn, embellish, enrich, deck, lard$_2$, festoon, trim$_2$, vest$_2$, emblazon, endow, enrich, garnish, imbue$_2$, infuse$_2$, inlay$_2$, ornament, replenish, season, bandage$_2$, hang, string$_2$, drape$_2$, inject, shower

qualitative/esthetic (negative): dirty, litter$_2$, pollute, smother$_2$, soil, spot$_2$, stain, taint, blot$_2$, burden$_2$, clutter$_2$, deluge$_2$, entangle$_2$, infect, riddle$_2$, bombard, shroud$_2$, daub$_2$, smear, smudge, spatter$_2$, splatter$_2$

Note: Alternators are underlined; non-alternators are presented in plain text. Subscripts indicate the membership of a verb in two endstate subclasses (but nothing is implied by the pairing of subscripts to readings). The verbs in this analysis are from Rappaport and Levin (1985) (see the Appendix).

the spatial distribution of the content as it changes location, and the endstate verbs *must* specify either the geometry of (the relevant portion of) the reference object as it changes physical state (becoming holistically affected) or the purpose of the action in terms of the abstract change of state of the reference object. Thus, we are making a claim about some of the elements of meaning that are conflated into locative verbs; however, these subclasses are not meant to exhaust the meanings of the verbs they contain.

As shown in Table 32, the verbs which accept the content-locative form may be rather transparently organized into seven clusters according to the spatial distribution of the content as it changes location. The verbs of three of these subclasses specify the "unit" of content that is transferred: either in single particles (e.g, *dribble*) or blobs (e.g., *slosh*), in an array of particles or blobs (e.g., *spatter*, *shower*), or in a continuous stream (e.g., *pour*, *inject*). (In this last subclass we included *flood*, as in the unaccusative intransitive *water flooded into the basement*; this is in keeping with our assumptions about lexicosyntactic representation.) Four of the other manner subclasses specify a change in the "internal" distribution of the content during the course of the action: a (usually) semi-solid mass is forced into a layered distribution against a surface (e.g., *smear*, *plaster*); a mass assumes the shape of a vertical mound (e.g., *pile*, *stack*); a mass becomes compressed or smaller in volume (e.g., *wad*, *crowd*); and a flexible object assumes a circular or coiled shape (around a reference object) (e.g., *twist*, *wrap*). The final subclass (i.e., suspension verbs such as *hang*, *string*, and *drape*) is itself a portion of a larger subclass of surface attachment verbs (including *nail*, *tape*, *glue*, etc.). None of these verbs really specifies the spatial distribution of the content *as it changes location*; rather, they specify the method of attachment or, in the case of the suspension verbs, the spatial arrangement of the *static* theme with respect to the reference object (e.g., *John hung the portrait in the East Room for years*). We have included the verbs of suspension in our analysis because they may be

extended easily into the content locative form, and even into the container-locative form (as we shall see).

One virtue of this analysis is that it succeeds by proposing only *one* dimension along which these verbs vary. We have resisted the temptation to propose other dimensions on the grounds that they either weren't inclusive enough (for all of the verbs which take the content-locative form) or were too inclusive (potentially applying to verbs which could never take the content-locative form). For example, one possibility which we haven't explored here is that a separate dimension of verb meaning, having to do with the *inception* of a transfer of content, is relevant to the manner thematic core. This dimension would cross-cut the spatial-distribution categories by clustering verbs according to whether an imparted force causes the ballistic motion of a mass (e.g., *spray, splash, squirt, inject*) or gravity causes the motion of a mass (e.g., *dribble, drip, pour, dump*) (see Pinker, 1989). *Inception*, in this sense, would also provide a more reasonable categorization of *spill, slosh*, and *slop*, which seem to care less about the spatial distribution of the content than about the accidental inception or careless execution of the action. Although such a dimension truly captures new information about the meanings of these verbs, it seems to do little additional work in providing the child with a sufficient reason for extending the content locative to a verb (whereas it may be relevant to other PASs; we argue below that locative verbs of forceful surface contact may be expressed in the *container*-locative form).

On the other hand, it is perfectly acceptable, from a learnability point of view, for there to exist a few verbs whose ability to take a PAS is not predictable from the meaning of the verb. Only *one* verb which accepts the content-locative form utterly failed to fall into our manner subclasses: *load*. (It may be possible, however, to formulate a manner subclass on the basis of the form-fitting relation of content and container in *load*.) Such a verb would constitute a

positive exception to the rule that verbs expressed in the content-locative form specify *some* manner in which the content changes location. Crucially, children could always learn the content-locative syntax of *load* through positive input. But the good news, from a learnability point of view, is that none of the manner subclasses applies to any verb which clearly *cannot* accept the content-locative form. Thus, none of these subclasses accept *fill* or *drench* or *adorn* as a potential member (unless the verb's meaning is misunderstood); the child would have no choice but to conclude, *eventually*, that these verbs are incompatible with the manner thematic core and its corresponding PAS (i.e., 15a).

For verbs which may be expressed in the container-locative form, Table 32 breaks down the twelve endstate subclasses into two dimensions: eight subclasses vary in the geometry of (the relevant portion of) the reference object as it changes state (becoming holistically affected); four subclasses vary in the purpose of the action, as specified in terms of the non-spatial properties of the reference object. Interestingly, for those verbs which specify the same type of reference-object geometry (e.g., a surface), the nature of the holism may vary from cluster to cluster. Three subclasses involve a surface: a surface may be completely covered by a layer of either solid (e.g., *tile*), semi-solid (e.g., *spread*), liquid (e.g., *flood*), or fine-grained particulate (e.g., *dust*) matter; a set of objects may be distributed over a surface, where the kind of objects is usually specified by the verb (e.g., *litter*, *spot*, *sprinkle*, *sow*); or a surface may (exhaustively) support the weight of the content (e.g., *burden*, *heap*). Two subclasses involve a medium: a medium may become layered by the content as a result of an action (e.g., *intersperse*, *vein*); or a medium may become co-extensive with the content as a result of the action (e.g., *drench*, *saturate*). Finally, three subclasses specify unique geometries: an edge may be covered by or attached to (at every point) a line of content (e.g., *edge*, *trim*); a container may be filled by the content (e.g., *fill*, *stuff*); and a path associated with a reference object may become blocked, preventing the movement of air (e.g., *choke*, *smother*), water (e.g., *dam*),

objects (e.g., *bind, lash*), and so forth. In all of these cases, we argue that the relevant geometry of the reference object, as well as the type of (potential) holism, is specified by the meaning of the verb. (We must speak of *potential* because, as we have seen, alternators in content-locative forms may not necessarily imply the affectedness of the reference object.)

The dimension of geometry, as outlined here, succeeds in clustering many, but not all, of the verbs which may be expressed in the container-locative form. The remainder of the endstate verbs, as well as some of those discussed above, appear to specify something about the *purpose* of the action; that is, it is possible for a locative verb to specify the non-physical change of state that a reference object undergoes by means of content being added to that reference object. *Non-physical change of state*, in this context, may be illustrated in four subclasses. First, verbs such as *stock, load*, and *pack* specify that the function of a container is realized (e.g., *John stocked the shelf with groceries*) or potentially realized (e.g., *John loaded the gun with bullets*; *John packed the bag with clothes*). Notice that these verbs specify more than just the filling of a container; thus, ?*John loaded the box with toys* sounds odd unless the box is thought of as a vehicle of some kind. Second, verbs such as *dab, squirt, spray*, and *splash* specify the forceful contact of a surface with some content. One reason for thinking that forceful surface contact is incorporated into the meanings of these verbs is their ability to accept forms in which the reference object is the oblique object of *at*: *John sprayed/splashed/squirted water at the boy*. It is even possible to demote or entirely eliminate the content NP with some of these verbs, as in the conatives *John dabbed at his eyes* (*with the handkerchief*) or *John splashed at the boy* (*with water*). These forms should be contrasted with the corresponding "gravity" verb forms: ?*John dumped water at the car*; ?*John poured water at the pitcher*; **John dumped at the car* (*with water*); **John poured at the pitcher* (*with water*). These contrasts highlight a component of meaning shared by all verbs of surface contact: the use of the *at*-phrase signifies the intention of the agent to contact a surface, whether by splashing it or by

hitting it (cf. *John hit at the wall (with his fist)*. Furthermore, it is quite plausible to assume that children are aware of this salient component of meaning. In fact, several children in Experiment 1 produced conative sentences with *splash*.

The remaining two purposive subclasses both involve some change in the quality or appearance of the reference object. We have divided up the verbs specifying such a change into two main clusters, *positive* and *negative*. These clusters could probably be further subdivided on the basis of whether the change is esthetic or more permanent/substantial. Examples include: *adorn, embellish* (positive esthetic); *endow, bandage* (positive qualitative); *dirty, stain* (negative esthetic); *bombard, infect* (negative qualitative). We note that verbs of suspension (e.g., *hang, drape, string*) may be extended into the positive esthetic cluster (e.g., *John hung the room with pictures*; *John draped the window with silk*). Although the categorization of some verbs as positive or negative may be difficult in a few cases (e.g. *inject*, depending upon how you take your medicine!), these two poles seem clear in general.

By proposing that purpose is compatible with the endstate thematic core, especially the purpose of changing the quality or appearance of an object, we have succeeded in clustering the remaining verbs which may accept the container-locative form. But why are *two* dimensions relevant to the expression of container locatives? If we look upon a locative event as a potentially complex causal chain--a series of causes and effects--then there is a sense in which the same physical event in space can be interpreted in terms of successively abstract goals: changing the location of the content; changing the physical state of the container; changing the the container in quality, appearance, or potential to function. In fact, the endstate verbs in Table 32 may be placed into four categories along a continuum of polysemy, as to whether they ordinarily specify the concrete change in physical state of a container, but nothing more abstract (e.g., *cover*); the concrete change in physical state of a container, plus something more abstract

(e.g., *drape*); the abstract change of state of a container, but plus something more concrete (e.g., *inlay*); the abstract change of state of a container, but nothing more concrete (e.g., *embellish*). Furthermore, this continuum is probably better thought of as a progression with a fixed directionality to it (i.e., cause-to-effect), a directionality which appears to be reflected in the extension and perhaps etymology of some of these verbs. Considerations such as these argue against any notion that we have introduced the "purpose" dimension merely as a convenient escape-hatch.

The analysis of endstate verbs according to reference-object geometry and purpose succeeds in clustering all of the verbs which may accept the container-locative form. If this analysis is to provide a potential solution to Baker's Paradox, however, it must also guarantee that none of the verbs which accept *only* the content-locative form fall into endstate classes. In this regard, let us focus on the circle/coil verbs (e.g., *coil*, *spin*, *wind*; we'll return to the case of *wrap*): shouldn't they alternate? After all, these verbs appear to specify *some* sort of reference-object geometry, just as the vertical-mound verbs (which *do* alternate) specify that the reference object must be a surface. Let us assume, for the sake of argument, that these verbs require a linear reference object. Under this assumption, then, why doesn't there exist an endstate subclass in English according to which a linear reference object becomes completely encircled? Why can't we say *John coiled the post with rope* if the rope completely encircles the post? Indeed, there seems to be no principled reason for this gap, and we would expect that some languages might treat such verbs as alternators. But English, for whatever reason, does not. This arbitrary fact about English highlights two important points. First, we take this opportunity to remind the reader that the linking rules, in our conception, don't care about what entities are affected in the world, only about what entities are taken to be affected *in the meaning of a verb*. Second, whether or not an entity is affected in the meaning of a verb appears to depend crucially on the membership of that verb in a relevant subclass. In this case,

specifying just *any* endstate, or even just *any* reference-object geometry (e.g., linearity), is not sufficient, despite the fact that reference-object geometry seems to be relevant in general. For this reason, it is clear that the relevance of a subclass to a given thematic core/PAS must sometimes, if not always, be learned on the basis of positive evidence (e.g., hearing verbs of surface-support, but not "linear-object encircling", in the container-locative form).

The final focus of our analysis will be on alternating verbs, which are underlined in Table 32. One of the goals of our analysis was to insure that each of the alternators appears in at least one of the manner subclasses *and* at least one of the endstate subclasses. In general, we have achieved this goal, with the positive exception of *load* on the manner dimension. In the exposition of subclasses above, however, we equivocated on the issue of precisely *how* alternating verbs specify the affectedness of contents and reference objects. In returning to this issue, let's begin with the following observation: the membership of some of the alternators in endstate subclasses appears to be unpredictable solely from the content-locative form of the verb. In these cases, the actual holism of the reference object or the actual purpose of acting upon the reference object may be apparent *only in the container-locative form.* Notice that this is a statement about the ability of children to predict syntactic privileges, not (in our view) about the meanings of verbs. That is, because the two locative thematic cores necessitate the affectedness of different entities, an alternator must at least specify the *potential*, but not necessarily the *actual*, affectedness of both the container and content in a particular way. In this context, the semantic shift known as the "holistic effect" accompanies the locative alternation in just those cases where an alternator specifies the potential, but not actual, affectedness of the container.

For example, how would a child be able to predict that *sprinkle*, but not *scatter*, alternates upon hearing each of them in a content-locative form? The problem here is that *sprinkle* in the

content-locative form specifies the potential affectedness of the reference object, a potential which may difficult for the child to identify. We know that this potential is incorporated into the meaning of the verb because *sprinkle* doesn't just specify *any* affectedness of the container, it specifies the type of potential holism and the geometry of the reference object (i.e., it belongs to the subclass of *surface distribution*, which has been independently established by the membership of non-alternators.) In particular, *sprinkle* implies a somewhat controlled, local dispersal of elements--over food, for example; *scatter*, on the other hand, implies more of a random and widespread dispersal, applying to such things as leaves and crowds. The affectedness of a sprinkled surface, but not a ?scattered surface, may thus be a function of subtle differences in verb meaning, inextricably linked with idiosyncratic facts about what sorts of entities are commonly taken to correspond to the arguments of a verb. But the question is: how are these differences in verb meaning discernible from a few utterances (and contexts) such as *John sprinkled peanuts on the ice cream* and *the wind scattered leaves over the yard*?

Our answer must be that the potential surface-distribution holism of *sprinkle* may be *un*predictable, depending upon the vagaries of input. This fact about the learning of *sprinkle* and presumably other alternators forces us, in the interests of learnability, to attribute a certain amount of conservatism to the child in such cases: in every case where the subclass of an alternator is unpredictable on the basis of prior exposure, our account predicts that the child will simply *not* extend syntactic privileges productively to that verb. Our guess is that the number of such cases is small. Most alternators specify that both the content and the reference object are actually affected in a particular way (i.e., the potential for affectedness is highly predictable; e.g., *stuff*); we would expect children to be able to predict that both locative PASs are acceptable for these verbs. Also in this category are verbs such as *spray*, in which one primary endstate reading is probably predictable (forceful surface contact) although other readings may not be (surface coverage or distribution, as in *John sprayed the wall with paint*).

To summarize our analysis: the subclasses pick out components of verb meaning which are *specific* to either the manner or the endstate thematic core. Of course, these components aren't *reducible* to their respective thematic cores--they refer to the (potential) affectedness of an entity *plus* a particular manner or endstate. But because the subclasses are formulated as specific to thematic cores, the analysis accomplishes the most important work of disallowing the irremediable overgeneration of non-alternators in ungrammatical forms, while still allowing for a substantial amount of productivity. On this account, certain facts must be learned conservatively: the relevance of a dimension, and of particular subclasses, to a thematic core; and the syntactic privileges of positive exceptions, which exist either because no subclass exists for the interpretation of a verb (e.g., *load*) or because membership of an alternator in a subclass is otherwise unpredictable (e.g., *sprinkle*). On the other hand, this analysis grants children the productive capacity to predict the appropriate syntactic privileges of non-alternators, in those cases where the subclass membership can be ascertained without exposure to a locative form, and of alternators, in those cases where the potential affectedness of an entity is highly predictable.

Finally, in cases where the miscategorization of a non-alternator has syntactic consequences (e.g., as when *fill* is misconstrued as specifying the distribution of content in a continuous stream), this account predicts that the syntactic error will be unlearned as the child revises his or her interpretation of the verb's meaning (on the basis of more exposure to contexts of the verb's usage). We may summarize the full sequence in such a case as follows:

1. Children learn, conservatively, the verb domains in their language, and the predicate argument-structures (PASs) that are available in those domains.
2. They use universal linking rules, plus domain-specific PASs, to "build" or triangulate on the thematic cores for a domain.

3. They use these thematic cores to organize the domain: to search for dimensions of verb meaning, and subclasses along those dimensions, which specify the particular way in which a thematic core may be realized in their language.
4. They learn, conservatively, the dimensions/subclasses of verb meaning which are relevant to the (high-frequency) verbs heard in a PAS.
5. Manner dimensions/subclasses are fixed before Endstate dimensions/subclasses (all things being equal) due to a general cognitive/perceptual "Manner Bias."
6. Verbs in a domain are classified with respect to learned dimensions/subclasses. *Fill* is miscategorized into a manner subclass. (Note: the data from Experiment 2 can't really distinguish between the case in which the manner and endstate of *fill* are independent (which may be the case for young children) or mutually constraining (like *stuff*). Indeed, if independent, there needn't be any endstate meaning at all; no endstate subclass or dimension may have been found for the verb, due to the manner bias, despite its being heard in the container-locative form (#4 above). Thus, the container-locative form of *fill* may persist in the lexicons of children because of its positive input, and *perhaps* also because of a fixed (independent or interpredictive) endstate meaning.)
7. Syntactic privileges are granted to verbs on the basis of dimension/subclass membership.
8. Misclassified verbs are eventually reinterpreted, especially as the manner bias is overcome--and abstract changes of state become more ingrained in cognition and language.
9. Syntactic privileges are "revoked" in cases such as *fill* as the erroneous component of meaning is dropped.

Is Each Cluster Specific to a Thematic Core?

The other side of the learnability question, with respect to the alternators, is the following: if we assume that children can always predict the ability of an alternator to accept both PASs solely on the basis of the verb's meaning, is it possible to formulate subclasses for which all of the members alternate (versus other subclasses for which none of the members alternate)? Subclasses, in this formulation, would be completely independent of, and thus not necessarily specific to, particular thematic cores (e.g., Pinker, 1989). One seemingly strong piece of evidence in favor of such an approach is the extent to which the alternators in Table 32 actually clustered together; that is, for many of the subclasses, either most of the members are non-alternators or most of the members are alternators (i.e., array of particles/blobs, layer, vertical mound, compressed mass, suspension, container, surface (support), function, and forceful surface contact). Although our subclasses were not designed to isolate alternators from non-alternators, clusters of verbs which share semantic, syntactic, and (in many cases) phonological properties emerge from the analysis. Furthermore, the manner and endstate interpretations of verbs such as *cram*, *crowd*, *jam*, and *stuff* seem to be mutually constraining, as we mentioned in Experiment 1; in these cases, "a mass is forced into a container against the limits of its capacity" (Pinker, 1989, p. 129). Why, then, should we try to tease apart the properties of alternating verbs (into manner and endstate components) which may be better left unanalyzed and independent of affectedness per se?

The best answer to this question is simply that the clusters (probably, no matter how they are formulated) are not perfectly predictive. This is important because in a system where the ability of a verb to alternate is based on its membership in a single subclass, the existence of true non-alternators in that subclass would lead to negative exceptions. For example, the verb *wad* is presumably part of the cluster that includes *cram*, *crowd*, *jam*, and *stuff*. However, *wad*

does not appear to alternate: *John wadded the paper into the hole* may be acceptable, but ?*John wadded the hole with paper* is odd. (On our analysis, the verb *wad* specifies that a content is compressed into a small mass or ball, as in *John wadded the paper into a ball*, but the potential involvement of a container is clearly secondary.) Other likely negative exceptions include *scatter* (similar to *spatter*, *splatter*, *sprinkle*, etc., but ?*John scattered the field with seeds*) and *stock* (similar to *load* and *pack*, but ?*John stocked the cans onto the shelf*; this marginal locative form is distinct from the use of *stock* to mean *store* in a non-motional sense, as in *John stocked the cans on the shelf*.) Of course, one may argue that these verbs aren't *really* locatives, but the fact remains that they bear a striking similarity to verbs which *are* locatives. Such similarity, coupled with the tendency to formulate subclasses of alternators, might lead children to produce ungrammatical forms. Crucially, the ability of children to *un*learn these errors would be hampered by an analysis of verb meaning which collapsed distinctions between manner and endstate.

On the other hand, by formulating manner and endstate subclasses, the potential for negative exceptions is avoided. In addition to accounting for the fact that some alternators and non-alternators may share components, it also accounts for two other observations. First, alternators in a cluster may actually agree on one dimension but differ in the other. For example, *jam* and *stuff* both specify a compressed-mass manner, but the particular endstate of *jam* is more applicable to path-blocking than to container-filling (e.g., *John ?jammed the box with clothes / stuffed the box with clothes; John jammed the sink with onion skins / ?stuffed the sink with onion skins*). Second, endstate verbs (alternators or non-alternators) may belong to more than one endstate subclass. Numerous examples of such polysemy are illustrated in Table 32. Both of these observations support a more "componential" approach, in which dimensions are specific to particular thematic cores.

Finally, we must ask the question: why do the clusters exist? We can offer two (mutually consistent) explanations for this phenomenon. First, the clustering of verbs, especially the phonological similarity of verbs in clusters, may be due to historical processes--independent of syntax-semantics linking. Mistakes or guesses about verbs, based on semantic or phonological analogy, may have been perpetuated generation after generation for their functional value (see Fodor, 1985). Second, the mutually-constraining nature of the manner and endstate components of verb meaning (i.e., their *inter-predictability*; Pinker, 1989) may be plausibly hypothesized to be an independent constraint on verb meaning--a species of a general constraint of coherence that applies to alternators. This would explain why alternating verbs that are similar in manner are also similar in endstate, and vice versa. That is, this constraint explains why alternating verbs *appear* to cluster in subclasses that are non-specific to individual thematic cores.

Developmental Evidence Concerning Subclasses

The ease, if not transparency, with which locative verbs may be classified into narrow subclasses provides some support for the psychological reality of the subclasses. Furthermore, the subclasses involve semantic elements of the sort that have been postulated independently to be conflated into the meanings of verbs--especially, the notions of force (Talmy, 1988), geometry (Talmy, 1983; Jackendoff, 1983), and the count/mass distinction. Yet we have little direct evidence that narrow subclasses actually play a functional role in language acquisition.

The evidence from Experiments 1 and 2 is certainly consistent with our hypothesis that children can predict the syntactic privileges of verbs on the basis of their membership in manner or endstate subclasses. In those experiments, we found that children are more than willing to produce the *fill*-content form, and our tests of association suggest that the misinterpretation of filling in terms of a *particular* manner is a primary source of the

overgeneration. In these experiments, we focused on one particular misinterpretation of *fill*: we assumed that a child might take *fill* to specify a pouring manner, because children are presumably exposed to many events which are both pouring and filling (i.e., pouring being a common means of filling). But were children who made syntactic and semantic errors with *fill* actually miscategorizing *fill* in a manner *subclass* having to do with the distribution of liquid in a continuous stream? This is impossible to answer in retrospect, of course, but the fact that children are delayed both in learning the particular endstate of filling, and in *mis*learning the particular manner of filling, is consistent with the view that before a child fixes the meaning of a verb in terms of a particular manner or endstate, he or she must have learned the relevant dimensions and perhaps subclasses for the locative domain.

The evidence from Experiment 6 is a bit more direct: children, as well as adults, can use models in order to learn the syntactic privileges of new verbs. Specifically, the novel holistic verb was interpreted either like *put* or like *cover* (with the consequence that relatively more content or container locatives were produced, respectively). Of course, it is a long jump from the forced choice between two disparate verbs such as *put* and *cover* to the fine-grained semantic distinctions made in Table 32. Thus, the results of Experiment 6 show only that *some* metric of similarity can be used in learning new verbs; future work must focus on whether, and how, that metric of similarity is defined in terms of a child's current lexical knowledge (e.g., one possible scenario is that children begin with a few verbs in relatively coarse-grained classes, which are later subdivided as more and more verbs are learned.)

Conclusion

In conclusion, we have presented a strong, multifaceted case for the hypothesis that a universal linking rule of Object Affectedness is used by children to predict the syntactic privileges of verbs, but that children must learn what counts as affected. Our experimental

evidence has shown that a causal connection between verb meaning and syntax, plausibly involving Object Affectedness, can account for the semantic and syntactic errors that children make in learning locative verbs, and for the ability of children and adults to predict the syntax of novel verbs on the basis of verb meaning. Our survey of the cross-linguistic literature supports the view that the affectedness of direct objects is a universal tendency, applying across domains in English and in other languages. The combined experimental and cross-linguistic evidence, especially in the light of learnability considerations, argue against any view in which the linking regularities are wholly learned.

In developing the view that the linking rule of Object Affectedness is innate, we found that two sources of mediation are relevant to the use of the linking rule: the operation of a *set* of linking rules in a domain and the clustering of the verbs of a domain into subclasses. The influence of these factors forced us to recast the linking rule in terms of the PASs of a domain and their corresponding semantic representations, or thematic cores. By framing the linking regularities at this level, rather than at the level of individual grammatical and semantic relations, we were able to state the rule more strongly: the matching of a verb to a thematic core is a necessary condition for the verb to take the corresponding PAS. However, a consideration of Baker's Paradox shows that determining the affectedness of an entity in the meaning of a verb cannot be *sufficient* for predicting its syntactic privileges. One way out of this paradox, following the lead of Pinker (1989), is to posit that children determine the sufficiency of a verb to accept a particular PAS (within a domain) on the basis of its membership in a semantically coherent subclass. In this regard, we argued that the pair of thematic cores corresponding to alternative PASs in a domain may structure the domain into two divisions (e.g., manner and endstate), and thus may guide the child's search for relevant subclasses. Finally, we demonstrated that an analysis of locative verbs into narrow subclasses is successful in disallowing the irremediable overgeneration of non-alternators in ungrammatical forms, while

still allowing for a substantial amount of productivity. It is this productivity, and any contraints upon it, which we must characterize if we are to understand language acquisition, innovation, and change.

References

Anderson, S. R. (1971) On the role of deep structure in semantic interpretation. *Foundations of Language*, 6, 197-219.

Aronoff, M. (1976) *Word formation in generative grammar*. Cambridge, Mass.: MIT Press.

Baker, C. L. (1979) Syntactic theory and the projection problem. *Linguistic Inquiry*, *10*, 533-581.

Bhat, D. N. S. (1977) Multiple Case Roles. *Lingua*, *42*, 365-377.

Boons, J.-P. (1974) Acceptability, Interpretation, and Knowledge of the World: Remarks on the verb PLANTER (to plant). *Cognition*, *2*, 183-211

Bowerman, M. (1982) Reorganizational processes in lexical and syntactic development. In E. Wanner and L. R. Gleitman, eds., *Language acquisition: The state of the art*. New York: Cambridge University Press.

Bowerman, M. (in press) Mapping thematic roles onto syntactic functions: Are children helped by innate "linking rules"? *Journal of Linguistics*.

Braine, M. D. S. (1971) On two types of models of the internalization of grammars. In D. I. Slobin,, ed., *The ontogenesis of grammar: A theoretical symposium*. New York: Academic Press.

Bresnan, J., ed. (1982) *The mental representation of grammatical relations*. Cambridge, Mass.: MIT Press.

Brown, R., and Hanlon, C. (1970) Derivational complexity and order of acquisition in child speech. In J. R. Hayes, ed., *Cognition and the development of language*. New York: Wiley.

Burzio, L. (1986) *Italian Syntax: A Government-Binding Approach*. Dordrecht, Reidel.

Carter, R. J. (1976a) Some constraints on possible words. *Semantikos* 1: 27-66.

Carter, R. J. (1976b) Some linking regularities. Unpublished manuscript, University of Paris VIII, Vincennes.

Chomsky, N. (1981) *Lectures on government and binding*. Dordrecht, Netherlands, Netherlands: Foris Publications.

Comrie, B. (1981) *Language Universals and Linguistic Typology*. Chicago: University of Chicago Press.

Crain, S., Thornton, R., and Murasugi, K. (1987) Capturing the evasive passive. Paper presented at the Twelfth Annual Boston University Conference on Language Development, Oct. 23-25.

Erteschik-Shir, N. (1979) Discourse constraints on dative movement. In T. Givon, ed., *Syntax and Semantics*. Vol. 12: *Discourse and Syntax*. New York: Academic Press

Fillmore, C. J. (1968) The case for case. In E. Bach and R. J. Harms, eds., *Universals in linguistic theory*. New York: Holt, Rinehart, and Winston.

Fodor, J. D. (1985) Why learn lexical rules? Paper presented at the Tenth Annual Boston University Conference on Language Development, Oct. 25-27. Written up as "The procedural solution to the projection problem," unpublished manuscript, City University of New York.

Fraser, B. (1971) A note on the spray paint cases. *Linguistic Inquiry* 2.4. 604-607.

Fukui, N., Miyagawa, S., and Tenny, C. (1985) Verb classes in English and Japanese: A case study in the interaction of syntax, morphology, and semantics. Lexicon Project Working Papers #3. Cambridge, Mass.: MIT Center for Cognitive Science.

Gentner, D. (1978) On relational meaning: The acquisition of verb meaning. *Child Development* 49· 988-998.

Gentner, D. (1982) Why nouns are learned before verbs: Linguistic relativity vs. natural partitioning. In S. A. Kuczaj II,, ed., *Language development*. Vol. 2: *Language, thought, and culture*. Hillsdale, N.J.: Erlbaum.

Green, G. M. (1974) *Semantics and syntactic regularity*. Bloomington: Indiana University Press.

Gropen, J., Pinker, S., Hollander, M., Goldberg, R., and Wilson, R. (1989) The learnability and acquisition of the dative alternation in English. *Language, 65*.

Guerssel, M. (1986) On Berber verbs of change: A study of transitivity alternations. Lexicon Project Working Papers #9. Cambridge, Mass.: MIT Center for Cognitive Science.

Hale, K. (1973) Person marking in Walbiri. In S. R. Anderson and P. Kiparsky, eds., *A festschrift for Morris Halle*. New York: Holt, Rinehart, and Winston.

Hale, K., and Laughren, M. (1983) The structures of verbal entries: preface to dictionary entries of verbs. Warlpiri Lexicon Project. Cambridge, Mass.: MIT.

Hirsh-Pasek, K., Treiman, R., and Schneiderman, M. (1984) Brown and Hanlon revisited: Mothers' sensitivity to ungrammatical forms. *Journal of Child Language* 11: 81-88.

Hopper, P. J., and Thompson, S. A. (1980) Transitivity in grammar and discourse. *Language* 56: 251-299.

Jackendoff, R. S. (1983) *Semantics and cognition*. Cambridge, Mass.: MIT Press.

Jackendoff, R. S. (1987) The status of thematic relations in linguistic theory. *Linguistic Inquiry* 18: 369-411.

Keenan, E. O. (1976) Towards a universal definition of "subject." In C. Li, ed., *Subject and topic*. New York: Academic Press.

Levin, B. (1985) Lexical semantics in review: An introduction. In B. Levin, ed., Lexical semantics in review. Lexicon Project Working Papers #1. Cambridge, Mass.: MIT Center for Cognitive Science.

Moravcsik, E. A. (1978) On the case marking of objects. In J. H. Greenberg et al., eds., *Universals of human language*, Vol. 3: *Word structure*. Stanford, Calif.: Stanford University Press.

Nwachukwa, P. A. (1987) The argument structure of Igbo verbs. Lexicon Project Working Papers #18. Cambridge, Mass.: MIT Center for Cognitive Science.

Ostler, N. D. M. (1980) A theory of case linking and agreement. Unpublished manuscript, Indiana University Linguistics Club, Bloomington.

Perlmutter, D. (1978) Impersonal passives and the unaccusative hypothesis. *Proceedings of the Berkeley Linguistics Society* 4: 157-189.

Pinker, S. (1984) *Language learnability and language development*. Cambridge, Mass.: Harvard University Press.

Pinker, S., Lebeaux, D. S., and Frost, L. A. (1987) Productivity and constraints in the acquisition of the passive. *Cognition* 26: 195-267.

Pinker, S. (1989) *Learnability and cognition: The acquisition of argument structure*. Cambridge, Mass.: MIT Press.

Rappaport, M., and Levin, B. (1985) A case study in lexical analysis: The locative alternation. Unpublished manuscript, MIT Center for Cognitive Science.

Rappaport, M., and Levin, B. (1986) What to do with theta-roles. Lexicon Project Working Papers #11. Cambridge, Mass.: MIT Center for Cognitive Science.

Schwartz-Norman, L. (1976) The grammar of 'content' and 'container'. *Journal of Linguistics* 12, 279-287.

Talmy, L. (1972) Semantic Structures in English and Atsugewi. Doctoral Dissertation, University of California, Berkeley.

Talmy, L. (1983) How language structures space. In H. Pick and L. Acredolo (Eds), *Spatial orientation: Theory, research, and application*. New York: Plenum.

Talmy, L. (1985) Lexicalization patterns: Semantic structure in lexical forms. In T. Shopen, ed., *Language typology and syntactic description*. Vol. III: *Grammatical categories and the lexicon*. New York: Cambridge University Press.

Talmy, L. (1988) Force dynamics in language and cognition. *Cognitive Science* 12: 49-100.

Tenny, C. (1988) The aspectual interface hypothesis. *Proceedings of the North East Linguistics Society* 18.

Vendler, Z. (1967) Verbs and times. *Philosophical Review* 66: 143-160.

Veyrenc, J. (1976) Sur la double diathèse d'object des enonces translocatifs. *Bulletin de la Société de Linguistique de Paris* 72: 241-273.

Wanner, E. and Gleitman, L. (1982) Language acquisition: The state of the state of the art. In E. Wanner and L. R. Gleitman, eds., *Language acquisition: The state of the art.* New York: Cambridge University Press.

Wexler, K. and Culicover, P. (1981) *Formal principles of language acquisition.* Cambridge, Mass: MIT Press.

Appendix

Syntactic Classes of Locative Verbs

LOCATIVE VERBS OF ADDITION

Content-Locative Only: dribble, drip, drizzle, ladle, pour, scatter, slop, slosh, spew, spill, coil, dump, shake, spin, twirl, twist, wad, whirl, wind

Container-Locative Only: adorn, blanket, block, bombard, choke, clog, cover, dam, edge, embellish, enrich, fill, deck, dirty, drench, festoon, inundate, lard, line, litter, plug, pollute, repopulate, saturate, shroud, smother, soil, spot, stain, stop up, stud, suffuse, taint, trim, vest, bandage, bind, blot, burden, clutter, coat, deluge, douse, emblazon, encrust, endow, enrich, entangle, face, garnish, imbue, impregnate, infect, infuse, inlay, interlace, interlard, interleave, intersperse, interweave, lash, ornament, pad, pave, plate, replenish, riddle, saturate, season, splotch, tile, dapple, ripple, vein, stock

Alternators: bestrew, dab, daub, hang, inject, pile, smear, spatter, spread, sprinkle, stack, string, cram, crowd, dust, flood, jam, shower, stuff, wrap, drape, load, pack, smudge, squirt, splash, splatter, spray, heap, plaster, slather, sow,

LOCATIVE VERBS OF REMOVAL

Content-Locative Only: delete, expel, grab, omit, recover, remove, seize, sever, steal, withhold

Container-Locative Only: absolve, acquit, balk, bereave, bilk, break (of a habit), cheat, cure, defraud, denude, deplete, depopulate, deprive, disabuse, disencumber, dispossess, divest, disarm, ease, exonerate, fleece, free, pardon, purge, ransack, relieve, rid, sap, unburden

Alternators: brush, iron, rub, rake, shovel, sweep, leech, bleach, comb, distill, dust, erase, expunge, filter, flush, hose, mop, prune, rinse, scrub, skim, shear, sponge, squeeze, strain, towel, trim, vacuum, wash, weed?, wipe, wring, cure, pluck, scrape, shave, unload, unpack, wear, clean, cleanse, empty, strip, bail, clear, drain, rob (possessional only), con (possessional only)

Note: The verbs in this analysis are from Rappaport and Levin (1985), with the exception of *smooth*, *brush*, and *wash*, for which no clear intuitions were forthcoming. In addition, our intuitions favor the treatment of *bestrew* as an alternator and *stock* as a non-alternator (see text for discussion).