Human Simulations of Vocabulary Learning*

Jane Gillette

Henry Gleitman

Lila Gleitman

Anne Lederer

Department of Psychology and

Institute for Research in Cognitive Science,

University of Pennsylvania

Running head: vocabulary learning

---

*We owe a great debt to the members of the CHEESE seminar who offered continuing theoretical insights and experimental suggestions as this work evolved, and we offer our heartfelt thanks to that group. We also had the benefit of informative discussions with Cynthia Fisher, Jerry Fodor, Dedre Gentner, Jane Grimshaw, Ray Jackendoff, Barbara Landau, and Jesse Snedeker. This work was supported in part by a grant to Lila Gleitman from the Marcia and Steven Roth Foundation, and in part by an NSF Center grant to the Institute for Research in Cognitive Science. Requests for reprints should be sent to L.R. Gleitman, Institute for Research in Cognitive Science, Univ. of Pa., 3401 Walnut St., Philadelphia, Pa. 19104. The e-mail address for the first author is gillette@cattell.psych.upenn.edu.

Abstract

The work reported here experimentally investigates a striking generalization about vocabulary acquisition: Noun learning is superior to verb learning in the earliest moments of child language development. The dominant explanation of this phenomenon in the literature invokes differing conceptual requirements for items in these lexical categories: Verbs are cognitively more complex than nouns and so their acquisition must await certain mental developments in the infant. In the present work, we investigate an alternative hypothesis; namely, that it is the information requirements of verb learning, not the conceptual requirements, that crucially determine the acquisition order. Efficient verb learning requires access to structural features of the exposure language and thus cannot take place until a scaffolding of noun knowledge enables the acquisition of clause-level syntax. More generally, we experimentally investigate the hypothesis that vocabulary acquisition takes place via an incremental constraint-satisfaction procedure that bootstraps itself into successively more sophisticated linguistic representations which, in turn, enable new kinds of vocabulary learning. If the experimental subjects were young children, it would be difficult to distinguish between this information-centered hypothesis and the conceptual change hypothesis. Therefore the experimental learners are adults. The items to be "acquired" in the experiments were the 24 most frequent nouns and 24 most frequent verbs from a sample of maternal speech to 18-24-month old infants. The various experiments ask about the kinds of information that will support identification of these words as they occur in mother-to-child discourse. In Experiment 1, subjects were required to identify the words from observing several extralinguistic contexts for their use (silent videos in which mothers are seen uttering the "mystery word" several times to the infants, with each such use cued by a beep or a nonsense word). The findings under these conditions mimicked the known learning trajectory for infants at the inception of speech and comprehension: Nouns are learned far more efficiently than verbs. Experiment 2 showed that the Experiment 1 results are best understood as concreteness differences that are correlated with lexical class membership in the common useage of mothers to young children. Experiment 3 presented (different) subject groups with 24 verbs under varying information Conditions; namely: (1) extralinguistic information; (2) noun-co-occurrence information; (3) both (1) and (2); (4) syntactic-frame information but with nouns and verbs represented by nonsense words; (5) both (2) and (4); (6) both (1) and (5). Each Condition led to greater identification success than the preceding Condition. Moreover, not only the number but the type of verb that was efficiently learned was different under the different information conditions. We discuss these results as consistent with the incremental construction of a highly lexicalized grammar by cognitively and pragmatically sophisticated human infants, but inconsistent with a procedure in which lexical acquisition is independent of and antecedent to syntax acquisition.

The rapidity of vocabulary learning by young children is notorious, with estimates in the range of 3 - 10 words acquired per day from about age 2 to 10 years (Carey, 1978). Here is how -- according to everyone who has ever considered this problem -- the children do it: They listen to the speech of adults, taking note of which words are said under which extralinguistic circumstances. For instance, they learn that “cat” is English for the notion 'cat' because this is the word that occurs most regularly in the presence of cats.[1] They learn that “give” is English for the notion 'give' because, across situations, this is the word that occurs most regularly in the presence of giving. That is, the learners pair the word to the world, parsing out of several encounters that which is common to its extralinguistic contexts. Just so.

Several questions are left open by this just-so story. Perhaps the hardest one concerns how to make good on a description of extralinguistic circumstances that will support the word-to-world pairing procedure. After all, there are so many different kinds of cats and cat-circumstances that we might wonder how the child navigates relevantly through them, managing to include the Manxes (tailless though they are), exclude the Boston Terriers (cat-faced though they are), and so forth. Worse, or so philosophers tell us, learners might conjure up absurd and endlessly differing representations for those entities we adults call "the cats." Worst of all, many relational words (including verbs and prepositions) do not describe the world directly; rather they describe some perspective on the world that the speaker has chosen. For example, to label the same event, a speaker might say "John gives the hat to Mary," or "Mary gets the hat from John." There is room to wonder, in light of sameness of the observations standardly supporting both these utterances, how a learner might decide that “give” means ‘give’ and not ‘get.’ (For a discussion of perspective verbs in this context, Gleitman, 1990). All in all, a moment's thought about the extralinguistic contexts for word use suggest that they are by no means uniquely or straightforwardly interpretable, even -- or especially! -- across instances.

Children are conceptually just like us

Very often these apparent problems are shrugged off in discussion of language learning by invoking the Principle of Charity. According to this Principle, young learners will correctly interpret the world in view and the adult’s speech intent just because child and adult are creatures of the same sort conceptually and motivationally, and so their conversations conspire to the same ends (for discussion, Spelke, 1985; Soja, Carey, and Spelke,1991; Landau, 1994; Pinker, 1984; Bruner, 1974/1975; and for some wonderful experimental evidence, Baldwin, 1991).

But if children conceptualize the world as adults do, another enigma in understanding their word learning immediately arises: The vocabulary learning functions for both production and comprehension differ across lexical class in ways that mismatch the input frequencies. Specifically, adults speak in simple but grammatical sentences to young children, thus using nouns, verbs, adjectives, prepositions, and so forth, much as in talk among adults (Newport, Gleitman, and Gleitman, 1977). Yet nouns heavily dominate the infant vocabulary (Bates, Dale, and Thal, 1995; Brown, 1973; Gentner, 1978; 1981; Goldin-Meadow, Seligman, and Gelman, 1976), a fact that appears to hold true under diverse linguistic and child-rearing circumstances (Caselli, Bates, Casadio, Fenson, Fenson, Sanderl, and Weir, 1995).[2] This effect is shown in Figure 1, from Bates et al, 1995, who give data for 1800 infants learning English: Verbs and function words are almost absent from the initial vocabulary, while nouns that name objects constitute almost half of it (most of the rest are routine words like "bye-bye," a heavy dose of animal sounds, plus proper nouns and a few spatial prepositions). This over-representation of nouns and under-representation of verbs, compared to their frequency in input speech, is true of child vocabularies up to about the third birthday. Some further piece of theoretical apparatus is needed to account for these input-output disparities.

Figure 1

Children are conceptually different from us

The dominant explanation for the early preponderance of nouns invokes changes in the child's conceptual structure over developmental time. Some of the caregivers' words just cannot be represented by the young listener so they pass through his ears without stirring up his brain. Indeed, the order of acquisition facts have been used by several authors as a tool for indexing and understanding conceptual growth (e.g., Huttenlocher, Smiley, and Ratner, 1983; Levine and Carey, 1982; McNamara, 1972; Mandler, 1992; Merriman and Tomasello, 1995; Nelson, 1981; Smiley and Huttenlocher, 1995). More specifically, Gentner (1978; 1981) proposed a conceptual explanation of why nouns are learned before verbs: Nouns typically describe objects, while verbs label the relations among those objects. In this sense, nouns seem simpler and hence more readily learnable by the least mentally sophisticated babies. More important, on this view the object representations that support noun learning would neces-sarily be in place before verb learning could properly begin.

The role of information change in word learning

To account for two aspects of word learning we have so far invoked two apparently contradictory notions: On the one hand, shared mental structure in child learner and adult tutor is said to facilitate learning because under these circumstances their conversational interactions can refer to the same reality. On the other hand, differences in conceptual structure have often been suggested as the explanation of mismatches between which words are most frequently heard and which of these are first learned by the very young child.

In the present work, we will consider another explanation for the order-of-acquisition facts; namely, that not all words are learnable from a single kind of input evidence. The required information for acquiring words from different lexical classes becomes available to the learner *seriatum,* not -- or not solely -- as a consequence of changes in conceptual status but rather as a consequence of solving prior parts of the language learning task. Specifically, it may be that only a small and limited stock of nouns can be identified solely from inspection of their standard *extra-linguistic contexts* of use (word-to-world pairing), while verb identification requires, in addition, inspection of their standard *linguistic contexts* of use (sentence-to-world pairing). If so, efficient verb learning would be delayed until the requisite linguistic representations are constructed (Landau and Gleitman, 1985; Gentner and Boroditsky, in press, for a related position). Lending initial plausibility to this information-centered approach to understanding the course of vocabulary acquisition, several studies have shown a close correlation between increase in the verb vocabulary and first indicators of syntactic knowledge toward the end of the second year of life (Bates et al, 1995; Gleitman and Wanner, 1982; Lenneberg, 1967).

To test this general idea, we investigated the psychological potency for word learning of several information sources for nouns versus verbs. One source of evidence for the meanings of these words, already discussed, is provided by the scenes and event streams that accompany their use; that is, nonlinguistic cross-situational observation. Other sources of evidence have to do with representations of the speech stream: the co-occurrence of semantically related words in sentences (e.g., the likelihood of food names with verbs like *eat*) and the syntactic structures in which words occur.[3]

To test for such effects of information change independent of conceptual change, we used Human College Sophomore as the experimental population. While we would not want to exaggerate the conceptual sophistication of these subjects, we can be quite confident of their competence and stability with respect to the ideas labelled by the words that are the stimuli in these experiments; namely, 24 nouns and 24 verbs that are among the most frequently encountered by the average English-learning child during the first two years of life.

The method of these studies was to present these simple words (masquerading as an audible beep or as a nonsense term such as *flurg* or *glorp)* for identification by adult Ss under varying informational circumstances. While to be sure, this method is artificial in various respects that we will point out, the stimulus materials themselves are realistic. They derive from actual videotape of mothers at play with their 18 to 24-month old language-learning offspring, and so represent the kind of talk in the kinds of setting in which infants begin to learn words.

In a useful sense, we believe, these studies are probes for the information value of aspects of the input -- the "stimulus," so to speak, that engages the acquisition engine. As such, this work is in the tradition of computer simulations in which, for example, it might be asked whether lexical classification could be achieved by a device that performs statistical analyses on the relative distribution of adjacent words in a corpus (e.g., Mintz, Newport, and Bever, 1995; Seidenberg and McClelland, 1989), whether a machine could learn the phrase structure of a language armed only with knowledge of word meanings (e.g., Grimshaw, 1981; Pinker, 1984), and so forth. Such investigations, like our own, ask whether the corpus representation supports learning under the proposed computational procedures. The goodness of the machine model is assessed by asking how faithfully it reproduces the target learning function (say, that of a 2-year old child). Here, we report <u>Human Simulations</u> instead of computer simulations by examining word identification in adults. If recent commentators are correct in their assessment of the categorial and pragmatic sophistication of young children in the relevant regards, there should be little difference between adult and child with regard to acquiring simple words so long as they are provided with the same information, for in both cases the problem reduces largely to one of mapping.[4] Insofar, however, as young children and adults bring very different conceptual apparatus into the task, their learning functions should look different too.

Part I: THE POWER AND SCOPE OF OBSERVATIONAL LEARNING

Experiment 1: Cross-Situational Observational Learning

This experiment has two goals: The first is to understand the kinds of words that are most efficiently acquired via observation of the ongoing scene. The second is to examine on the basis of the results the internal properties a successful cross-situational learning procedure must have. To assess how nouns and verbs occur in caretaker-to-child conversation in ways that bear on the vocabulary acquisition problem, we asked Ss to identify words just from observing several videotaped mother-child contexts without linguistic accompaniment. This was accomplished simply by turning off the audio.

Procedure:

Subjects were tested in small groups of two or three. They were shown silent video of mothers playing with their child. Their task was to identify the "mystery" noun or verb that the mother was actually uttering. The mystery word was represented by a beep that sounded just at the point on the videotape when the mother had actually uttered it. Subjects were asked for their current conjecture as to the identity of the mystery word after each beep, i.e., from accumulating cross-situational evidence. After hearing six beeps corresponding to six maternal utterances of the same word, the subject was invited to reconsider all the input and to offer a seventh (henceforth, *Final*) conjecture. For each subject, this procedure was repeated for 16 different words, half of them nouns and half verbs.

Materials and stimulus construction

Source of the materials: The sources of the videotaped conversations (each about one hour long) were four mothers and their toddlers, taped in their homes. The experimenter brought along a bag containing several new toys to capture the child's interest. The mother was asked to "play naturally" with her child, using the toys if and when convenient.

Choice and arrangement of test words: The frequency of each noun and each verb in the taped sessions was calculated from session transcripts, and the 24 most frequent nouns and 24 most frequent verbs in the sample (with the constraint that a chosen item had to have appeared in the sample from at least two mothers) were selected for investigation. Frequencies of the chosen ("target") words in the 4-hour corpus ranged from 7 to 119 (mean = 29 overall, 22.9 for nouns and 35.3 for verbs). With a few exceptions (the nouns *pilot* and *peg* and the verbs *hammer* and *pop)* these items are all extremely frequent both in adult-to-adult and adult-to-infant speech, and show up again and again among the lists of first nouns and verbs that children utter and understand.

For nouns and verbs separately, the 24 items were divided into three subgroups of the high, middle, and lower frequency items. The target nouns and verbs were randomly assigned to three stimulus lists with the restriction that each list contain equal numbers of items from each frequency group. Each subject saw only one of these lists, i.e., 8 verbs and 8 nouns per subject. The decision that individuals respond to only a third of the 48 items was made based on pilot studies which showed that otherwise the procedure would be unduly tedious . The noun-verb order of presentation of targets in each list was randomly generated and kept constant across all three groups. For half the subjects in each list-group, the presentation order was reversed.

Composing the videos: For each target word, six separate uses of the word were selected to compose a block of trials. For the lower frequency targets, all or most instances from the corpus were used, avoiding only instances where the mother was speaking of something outside the view of the camera. For higher frequency targets where there were many options, segments were selected on the following additional bases: (1) We avoided segments that had already been used for other target words in the same grammatical category. (2) We tried to draw instances from more than one mother/child pair. (3) We excluded segments where one might be able to read the mother's lips.

Once the particular target use had been chosen, a test segment was constructed: The segment began approximately 30 seconds before the occurrence of the target word and continued for about 10 seconds after the target was uttered (or slightly longer when necessary to avoid unnaturally cropping a coherent event). The beep was inserted exactly when the mother uttered the target word. A thirty second introduction was chosen because, according to pilot subjects, this was long enough to establish the simple gist of the situation while not so long as to bore them. In theory this would result in a 40 second video clip with one beep (which would be followed by five more 40 second clips). In natural speech to young children, however, mothers often repeat themselves. If the mother repeated the target word within less than 40 seconds, the next beep was inserted at that point rather than this second beep having its own 30-second introductory material. The tape segment that S now saw, always containing exactly six beeps, thus varied between 71 seconds (in case the mother kept repeating the word) to 273 seconds (in cases where there were no such repetitions). The mean length of the videos was 162.7 secs; (nouns 155 secs, verbs 170.4 secs). The number of separate video clips ranged from 1 to 6 (Mean 3.4 clips per target; 3.1 for nouns and 3.6 for verbs). The number of different mother-child pairs per target ranged from 1 to 4 (mean 2.3; 2.2 for nouns, 2.4 for verbs).

Subjects

Mother-child pairs: The children were 3 boys and a girl 18-24 months of age (mean 22.3 months) with MLU < 2.

Experimental Ss: The subjects were 84 undergraduates at the University of Pennsylvania, 50 men and 34 women. All were enrolled in an introductory psychology course and received extra credit for their participation. Subjects were randomly assigned to one of the three stimulus lists, and to one of the two orders of presentation.

Instructions

Subjects were told the truth: that we were interested in how well they could identify a word simply by observing the contexts in which it was uttered. They were instructed to write down their best guess of what word the mother was saying each time they heard a beep. Each target was identified to them in advance as either a noun or verb (If they had any trouble with this distinction, we gave them some examples). They were aware that the target was never uttered during these segments without their hearing a beep. They also knew that this target remained constant for all 6 beeps; however, they were encouraged to change their guess from one beep to the next if the new observed situation caused them to consider a different interpretation. After the sixth conjecture, they wrote down a seventh (*Final*) choice, which did not have to be the same as any of the first six.

Scoring

Responses were scored correct only if they were morphemically identical to the target (though differences in number, tense, etc. were ignored). Initially, we considered weaker criteria of success: After all, what's really important is not that subjects come up with exactly the mother's word, but whether they land in its general semantic neighborhood on the basis of the accumulating extralinguistic evidence. Assessing the nature and severity of errors in the interest of making this distinction turns out to be a rather complex matter, however. Consider for example possible misidentifications for the word elephant. These may include synonyms (e.g., pachyderm), the kind of "error" that one would like to discount in evaluating the results for it confirms rather than disconfirms S's ability to glean the meaning of the target from the information provided. At the other extreme are choices that indicate that the observer mistook the very referent the mother had in mind (e.g., ball or Mommy for the target elephant). A more problematic failure concerns the representation implied, e.g., responses that are superordinates (toy, puppet) and partitives (trunk) of elephants rather than the elephant itself. No simple similarity assessment will do for these kinds of response. A contextless measure for the similarity of meaning between puppet and elephant will show that they are semantically disparate, but we know that in the observed scene the elephant that the mother was naming was in fact a puppet. The point is that there is no simple and general way to assess how semantically close the subject is to the correct conjecture; all of the guesses are true of the scene "in some way."

We did, however, evaluate an alternative scoring that gave credit when the nontarget-word response reflected knowledge of what physical object or event the mother was referring to. To do so, we showed three raters each video and told them the actual word the mother had uttered in the scene. The raters then judged whether in each S's Final choice for that item he or she had the correct referent in mind, whatever the actual label. The raters were told to be extremely liberal in this assessment. Thus puppet and even trunk were to count as revealing of the right referent for elephant. For verbs, if the target was give, the raters were to score Ss correct if they responded with any motion verb that referred to physical possession or transfer of some object (e.g., hold or grab). Average agreement among the three raters was 84%. This extravagantly liberal scoring method raised scores slightly. However, because it did not alter the structure of the findings -- and is questionable in the first place -- we report all findings according to the exact target method.[5]

Findings

Control variables

In a repeated measures ANOVA on the 48 items, there were no within item order effects (F=1.46, df=1,46 p <.24), so findings for the two orders of presentation are collapsed in the analyses that follow. We also collapsed across corpus frequency (the number of times that each of the test words appeared in the four maternal samples), for a linear regression showed no effects of this variable either (F = 1.12, df = 1,46 p < .3). This is no surprise because even though the test words varied in frequency among themselves, all were among the 24 most frequent nouns and verbs in our corpus and almost all are highly frequent in maternal usage.

Identification as a function of Lexical class

The most dramatic result of these investigations, as we will now document, is the difference in how amenable nouns and verbs are to identification solely on the basis of non-linguistic observation.

Subjects' selections were drawn from a surprisingly narrow pool given the unconstrained nature of this task. On the Final choice, only 116 different nouns (including all 24 targets) and only 66 different verbs (including 17 of the 24 targets) were offered by the 28 individuals who saw videos for each.[6] But despite the larger choice set considered for the nouns, the percentage of correct choices was much higher for nouns than for verbs. Table 1 shows the percent correct, on the Final choice, for each of the nouns and verbs. The gross effect of word type is very obvious: 45% of the nouns but only 15% of the verbs were identified correctly on the Final trial. This lexical class difference was highly significant whether tested over items with ANOVA (F = 12.57, df = 1,46 p < .001) or subjects (paired t-tests across all Ss yielding a t of 9.79, df = 83, p < .0001). There was, however, also a significant effect of lists. In one of the three lists the noun-verb difference was negligible as shown by a significant list-by-lexical class interaction in an ANOVA conducted across items (F = 3.54, df = 2,42 p < .04). Analyses over Ss gave the same general results. While paired t-tests over two of the lists were highly significant with t's of 34.46 and 9.61 respectively, the third list was not (t = 0.906, df = 27, ns).[7]

Table 1

As another indicator of the dramatic noun-verb difference, notice in Table 1 that 8 (1/3) of the verbs, namely, know, like, love, say, think, have, make, and pop were never correctly identified by any subject, whereas every noun target was identified by at least one subject. Moreover, removing these eight most-difficult items from the analysis raises the percent correct score for verbs to 23%, still only half of the 45% correctnesss score for the nouns ($F = 4.41$, $df = 1, 38$, $p < .05$).

Of course, absolute percent correct in these studies cannot be interpreted literally onto real world learning. After all, the taped videos don't take in the whole scene visually, there may be some prior context before the 30-second onset that might be relevant to constructing the conversational gist, children in any case could get 7 or 8 or 800 exposures before they learn a word meaning, some incidents in our maternal tapes might have been especially uninformative, etc. The potentially interpretable result is the massive <u>difference</u> between the noun and verb outcomes when extra-linguistic information must function in the absence of other cues to word meaning. The obtained noun-preponderance in a word-to-world pairing procedure by adults reproduces a property of word learning in presyntactic infants. This is a first suggestion that while the machinery of extralinguistic observation may be sufficient to account for the learning of first nouns, *taken by itself* it may be too weak for efficient verb learning, whatever the age or mentality of the learner.

<u>Victories for cross-situational observation</u>

As we will now show, despite the absolute differences, performance for both lexical classes improves across trials just as any theory of cross-situational learning would predict. Table 2 shows the most frequent selection on the Final trial. For example, on trial one for the target *nose,* 9 subjects conjectured *elephant* and no other single conjecture was as frequent. On the Final trial for this same target, 19 subjects, a majority, conjectured *nose*, a clear victory for cross-situational learning (For ease of inspection, the Table uses boldface whenever the most frequent choice was the correct one).

<u>Table 2</u>

As a measure of consistent interpretatin of the scenes, we also counted the number of different conjectures for each trial of each item, a measure we will call *scatter*. There was a wider scatter of conjectures for verbs (mean of 11 different conjectures per verb target) than for nouns (mean of 8.2; $F = 7.57$, $df = 1, 46$, $p < .01$). Perhaps surprisingly, for neither word class did scatter decrease across trials. This effect was shown quantitatively in a repeated measures ANOVA across trials 1 - 6 with a significant effect of lexical class between items ($F = 13.6$, $df = 1,46$, $p < .001$) but no effect of trial ($F = 0.78$, $df = 5,230$, $p < .60$) or trial X lexical class interaction ($F = 0.75$, $df = 5,230$, $p < .59$). Another measure showed, however, that fewer individuals were contributing to the scatter on the later trials than on the earlier ones, and that more of the subjects were arriving at the target. So as a group the subjects were profiting from accumulating evidence, especially for the nouns. Figure 2 shows this effect, measured by mean number of correct choices by trial. A repeated measures ANOVA across items demonstrates that performance improves significantly across trials ($F = 6.74$, $df = 5,230$ $p < .0001$).

Figure 2

Problems for cross-situational observation: Two failures of tiger-constancy

As we have just seen, though the absolute level of verb identification was pretty pathetic (only 15%, despite the simplicity and frequency of the items), verb as well as noun identification improved across trials (Figure 2). Thus in principle the machinery of cross-situational observational learning seems to be operating successfully for both of the lexical classes investigated, no matter that progress seems slow and highly errorful for the verbs.

However, there are two important provisos that provide further impetus for limiting the role of this procedure as a mechanism of verb-vocabulary acquisition. Both have to do with the requirement that experience, while it may be variable and temporarily misleading, must not yield convergence onto false choices. To concretize, let us term this requirement *tiger constancy*, defined as the greater tendency for tigers, in preference to any other object, to be in view (and/or to be "salient," "pertinent," etc.) when "tiger" is uttered. To be sure, a tiger will not be visible on each such occasion for one might be discussing, *inter alia*, future trips to the circus. But across situations of use, even if with plenty of bumps along the way, there ought to be a growing predominance of those for which 'tiger' commends itself as the semantic conjecture while other conjectures would receive only sporadic and unsystematic support; *mutatis mutandum.*

Tiger-constancy appears to be operating quite well for nouns in the current experiment: The most frequent choice is the correct target for 9 of 24 items on the first trial, rising to 15 by the Final trial (62.5%). For verbs, the most frequent guess is the target for 4 of 24 items on the first trial, rising to only 6 by the Final (25%). Restating, for 18 verbs not only was there not convergence by a plurality of subjects to the target, there was *convergence toward a false target*. If this experimental finding reflects real word learning in this regard, there is some motivation to reduce the burden that observation is to bear.

Examined more closely, the false-convergence problem seems to have an associated feature that would make recovery from error unlikely: Subjects conjecture more global categories in response to the increasing number of observations. In the present experimental setting, three false targets are very popular with the Ss, and each of these accounts for over 10% of all false conjectures. 14.8% of all false verb conjectures on the Final trial are <u>look</u> and 11.5% are <u>play;</u> 10.6% of all false noun conjectures on the Final trial are <u>toy</u>. More worrisome is that the incidence of these general terms increases by trial. Figure 3 graphs this rise for the noun <u>toy</u> and the verb <u>look</u>, the items in each lexical class that represent the extreme of this phenomenon. A repeated measures ANOVA shows that these curves indeed are rising significantly over trials ($F = 2.31$, df 5, 220, $p < .05$) and a significant trial by lexical class interaction shows that this tendency is increasing more rapidly for *look* than for *toy* ($F=2.78$, df=5,220, $p<.02$).[8]

<u>Figure 3</u>

A moment's reflection makes clear why this effect should be true of cross-situational observation, as it has always been construed (see Hume, 1738 and Pinker, 1984 for explicit versions): It is because the learner is hypothesized to *remember* the past and present contexts for the occurence of the novel word, and to parse out those properties of the scenarios that are applicable to them all. The hoped for outcome of such a procedure is that the subject will zoom in on some narrowly conceived target. But as we see here it is just as possible for observers to take the fatal step of increasing the generality of their chosen word, selecting one that covers just about everything. We will explore in Experiment 3 machinery that makes use of cross-situational observation as a significant cue to word meaning, but which is constrained by several linguistic cues that can correct for its defects as a stand-alone cue.[9]

. Item specific effects

All the analyses presented so far collapse across the members of each lexical class, and yield a substantial advantage for nouns over verbs. However, there are also large differences in success rate within the lexical classes. Though nouns are easier overall to identify from their contexts than verbs, still there are easy and hard nouns, and easy and hard verbs. The most casual look at Table 1, which presented percent correct on the final trial for each noun and verb, suggests that something like "concreteness" or "imageability" is accounting for many of the item effects within lexical class. For instance, looking at the nouns, it appears that "basic level" whole objects (e.g., ball) are the more readily identified from contextual inspection than are names (e.g., Mommy) and superordinates (toy).[10] For the verbs, those that referred to physical acts (e.g., push) were more often identified than those that referred to mental states (want).

This leads to the not-so-profound thought that the machinery of observation allows identification only of observables.[11] Of course, what is "observable" is not so easy to describe. The elephant, the puppet, and the toy on the floor are equally easy to see, i.e., they are the same physical object. But in another sense, among these three it is only the elephant that is "observable," -- the representation plausibly referred to in deictic gestures and utterances ("Look at the elephant" or "This is an elephant").

Our next task, then, was to ask whether the successes of our subjects in the contextual learning situation were confined to the identification of concrete things and events. It could be that the obtained differences between noun and verb identification in the present experiment are artifacts of a distinction of abstractness that differs systematically for the two lexical classes, at least for the 48 items that represent maternal word-preference in talking to novice learners.

Experiment 2: Imageability and Contextual Learning

Very young children's vocabularies are heavily loaded with "concrete" and "pictureable" words and thin on "abstract" ones. One factor that is often said to play a role here is *constraints on the learner's conceptual repertoire*, with physicalistic concepts somehow more cognitively accessible to young children than mentalistic ones. Alternatively or in addition, there may be *constraints on the learner's information base*, with physicalistic concepts the only ones that can be readily matched up with the scenes that accompany input utterances. That is, if a novice is restricted to a word-to-world pairing procedure, she may fail to learn abstract words no matter how sophisticated her representational capacity. The Human Simulation paradigm is especially useful with regard to disentangling these issues, for it rules out cognitive-developmental interpretations of failures to identify what the speaker is referring to in favor of information-based interpretations. For surely none of our college sophomores has conceptual trouble with the ideas coded by the 48 test items, nor do they avoid conjecturing nonbasic words.[12] Rather, it may be that one can only learn “pictureable words” if the sole source of evidence is (moving) pictures, as in Experiment 1. To find out, we carried out an assessment of this variable with a new group of subjects.

Procedure

Subjects rated the 48 frequent maternal words for imageability. An imageability (rather than concreteness) instruction was chosen because on the face of it this seems to reflect most simply the hypothetical distinction between words that can be learned by viewing their instances and those that cannot.

Materials

We prepared three lists of the test items. One contained all 48 words from Experiment 1 in alphabetical order, with "to" preceding the verbs and "a" preceding the nouns (this was to disambiguate the classification because many of these common words, e.g., ball, bag, hand, hammer, kiss have both noun and verb uses); this list also contained 20 adjectives and prepositions as filler items. The second list consisted of the 24 test nouns, in alphabetical order. The third was the 24 verbs only, in alphabetical order. In these last two cases "to" and "a" were not required because the instructions mentioned lexical class. Each word was followed by a labelled scale, as follows:

1	2	3	4	5	6	7
NOT AT ALL IMAGEABLE					EXTREMELY IMAGEABLE	

Instructions

Instructions appeared in written form on a face sheet. The task was to rate each word for imageability by circling one of the numbers on the 7-point scale: "Any word which, in your estimation, arouses a mental image (i.e., a mental picture, or sound, or other sensory experience) very quickly and easily should be given a HIGH imagery rating; any word that arouses a mental image with difficulty (or not at all) should be given a LOW imagery rating." Subjects were given as examples of high-imagery items the adjective sweet and the preposition above; as low imagery items, ambitious and of.

Subjects

Twenty undergraduates at the University of Pennsylvania, given lab credit in a Psychology course for their participation, responded to one of the three lists; thus the total number of subjects was 60.

Findings

Imageability:

Table 3 shows the imageability results for the separate and merged lists. Predictably, given a mother's reasonable usage to an infant, the items are heavily skewed to the imageable end of the scale, though there is plenty of variance among items.

Table 3

The most reasonable comparison of noun and verb imageability was within a list that merged them all (and added various fillers); that is, that tested the concreteness of all 48 words on a single scale. As inspection of Table 3 (columns 3 and 6) shows, imageability is almost perfectly correlated with the noun/verb distinction, with all the nouns except thing being rated as more imageable than any verb other than hammer (Pearson Correlation Coefficient .82 $p < .0001$).[13] However, the very strength and stability of this category-based outcome suggested that the merged list, by containing and perhaps emphasizing the noun-verb contrast, might have artifactually legislated a distinction that it set out to investigate. That is, the subjects might have explicitly compared the nouns with the verbs thus generating a response strategy that exaggerated differences in their imageability ratings. So we constructed new separate stimulus lists for nouns and verbs, with different subject groups providing the responses for each list. The results are shown in Columns 2 and 5 of Table 3. The range of verb ratings hardly changed between the merged and the separate lists. More importantly, the original lexical class effect reappeared: Imageability scores are higher overall for nouns (5.75) than for verbs (4.43; $t = 4.12$, $df = 46$, $p < .001$). The correlation of concreteness on the separate list with the noun/verb distinction drops to .52 but remains highly significant ($p < .0001$).

Imageability and identification by extralinguistic observation:

We now asked about the extent to which imageability predicts success in identifiability. We compared imageability scores in the present experiment with success rate in the identification task of Experiment 1, using multiple regression. Using the merged list, there was a significant effect of concreteness ($F = 18.86$, $df = 1,44$, $p < .0001$) but no effect of Lexical Class ($F = .03$, $df = 1,44$, ns) or Lexical Class x Concreteness interaction ($F = .42$, $df = 1,44$, ns). The same results are found for the separate noun and verb lists. Concreteness accounts for so much of the variability that the noun-verb distinction, if it has any effect at all, is not detectable (for related results, Gentner, 1981).

Finally, we asked whether imageability scores for nouns and verbs (Table 3, columns 2 and

5) predicted the item effects within lexical class that were obtained in Experiment 1. That is, were the items, *within* lexical class, that were least imageable just those that had been hard to identify from cross-situational observation? The answer is basically yes: There is a statistically reliable correlation between imageability and identifiability for verbs ($r = .43$, $p < .04$). For nouns, there is a trend in the predicted direction but it is not significant ($r = .04$, $p < .09$).

Discunssion

The results just presented suggest that, technically speaking, it is not a difference between nouns and verbs *per se* that accounts for the ease with which words are identified by inspecting the environments for their use. The account is a more mundane and ultimately tautological one, namely that only observables -- the most "pictureable" or "imageable" items -- can be efficiently acquired by observation operating alone. This shows up as a massive advantage for nouns over verbs in the early vocabulary of children, and in Experiment 1, just because of the particular selections that mothers make among items in these lexical classes. Presumably if the mothers said "thought" and "knowledge" as often as they say "think" and "know," we would find that frequent nouns, too, are hard to glean from inferences based on scene inspection alone. Notice particularly that the probabilistic nature of the noun-verb distinction in early child vocabularies (Table 1) is easier to understand if, as the present results suggest, they are artifacts of imageability: The child word learner at the earliest stages of language exposure is limited to the information provided by the observeable extralinguistic contingencies. If observation provides the sole information base, then nouns labelling concrete nominal categories should be easiest to acquire. Moreover, verbs like throw and come will be easier to acquire than want and know even if the learner has equal conceptual access to physical-action categories and mental-state categories.

Part II

LINGUISTIC SUPPORTS FOR VERB LEARNING

The experimental findings so far reported, contrived though they are, hint that observation is an inefficient and errorful basis for most word learning beyond the animal noises and concrete basic-level nominals. Applicability of these findings to child language acquisition gains plausibility by noticing their consistency with learning data in the literature. Caselli et al (1995) and Bates et al (1995; see again Figure 1) report that from ages 12 to 16 months vocabulary grows at the rate of about .4 words per day. At 18 months, the rate has already risen to 1.2 words per day, that is, a three-fold increase, and verbs and functors are beginning to make their appearance in speech and comprehension. Most relevant here, the correlation of these changes in vocabulary size and constitution with the appearance of multiword speech is so large as to suggest a cause-and-effect relationship (Lenneberg, 1967). We have hypothesized elsewhere (Fisher, Hall, Rakowitz, and Gleitman, 1994; Gleitman and Gleitman, 1996; Mintz and Gleitman, 1998) an incremental learning machinery with at least the following properties: It acquires a small stock of nouns by word-to-world pairing, and then uses that stock of nouns as a scaffold for constructing representations of the linguistic input that will support a more efficient learning procedure. If so, we ought to be able to reproduce the hypothesized effects of linguistic-contextual knowledge on word identification by revising the information made available in the Human Simulation paradigm.

Experiment 3: **Linguistic supports for verb learning**

The present sequence of studies concentrated attention on the verbs that were so difficult to identify via the scene evidence in Experiment 1. It was asked whether adult subjects do better when provided with alternative or additional cues beyond scene inspection. Each of six such experimental *Conditions* presents (different) subject groups with varying cues and cue combinations that are typically available to children in their everyday conversational interchanges.[14]

Procedure:

The procedure for each Condition of this experiment was the same, and much the same as in Experiment 1. Subjects were tested individually, and were asked to identify mystery verbs. As before, they always knew that the visual and/or linguistic contextual cues were derived from actual interactions of mothers with their infants. This time they were not asked to respond after each cue for a single word. They simply offered their conjecture after receiving the full set of cues for that word (this is equivalent to the *Final* conjecture of Experiment 1).

Materials

The source of the materials was a set of videotaped play sessions of about an hour each of 8 mothers and their young children (mean age = 20.9 months, mean MLU = 1.5) in a laboratory room at the University of Pennsylvania, culled from a large archive recorded several years earlier. As for Experiment 1, these sessions were transcribed and searched for the 24 most frequent verbs within and across mothers. These are listed in Table 5 in order of their frequency. The new verb set overlaps 66% with that of Experiment 1 and again are among the most frequent in English usage. The differences in the two sets of 24 verbs are attributable largely to the changed toys the experimenter brought into the situation.

For each verb, all the uses from 3 of the 8 mothers (chosen randomly for each verb) became the stimulus set. The least frequent item (pull) in the corpus of the 24 verbs so chosen occurred 4 times during the 3 sampled hours of taped material, and the most frequent (go) occurred 84 times, for a total of 349 sampled occurrences, across verbs. All 349 instances were presented in some form to (different) subject groups. Thus, for example, because call occurred seven times in the three sampled hours of taping, subjects in each of the experimental Conditions had seven call-stimuli to consider as the basis for their conjecture. So information per item was a direct function of its frequency of use by the mothers. This contrasts with the method of Experiment 1 which arbitrarily provided 6 stimuli per verb and noun item.[15] The six experimental Conditions and the materials they used are shown in Table 4.

Table 4

Condition 1: Cross-situational observation

This Condition replicated the silent video manipulation of Experiment 1 with the following changes: Two judges selected the videotaped stimuli for each exposure to each of the 24 verbs according to the following criteria. They were shown 2-minute long videotapes, audio turned off, during which the mother uttered one of the verbs at exactly the 1-minute mark. The judges were informed which verb it was that the mother had uttered. They were asked to inspect the full 2 minutes of silent video and to pick out the sub-segment that in their view "accounted for why the mother said that verb when she did." For example, one sentence was "Show me your truck." The judges selected the segment during which the child was holding the truck up for his mother's inspection. (Note then that these videoclips were in general much shorter -- 20 seconds on average -- than those in Experiment 1, where they had averaged 170 seconds).[16] Subjects in the present experiment would therefore see an act of showing just when the task was to guess show, whereas a subject in Experiment 1 would have heard the beep just when the mother said show, regardless of the fact that the showing act might have occurred well before or after this. The upshot is that the videos of Experiment 1 gave better gist information (they were much longer), but the videos of the current experimental Condition corrected for the poor timelock of verb utterance to physical act. When the two judges had independently made their segment choices, the results were combined so as to prepare a tape starting at the earliest point selected by either judge, and ending at the latest point selected by either of them. All the segments for each verb were then spliced together with a brief intervening pause between segments, and presented to S, one verb at a time. See Figure 4 for a schematic of this Condition, for the verb *call*.

Figure 4

Condition 2: Noun cooccurrence

In this Condition the subjects never saw the videotaped scenes. Instead they were presented with a written list of the nouns that occurred with the verb in the (same) maternal sentences used in the videotape. These nouns were listed in alphabetical order for all occurrences, by sentence. The alphabetical order, as the Ss were informed, was to mask any inference to syntactic structure from their actual serial order in the maternal utterance. This Condition models a situation in which the learner, in possession of knowledge of many nouns but not of language-specific syntax, notices that some of these known nouns occur frequently with a particular novel verb, providing a basis for inferring its meaning e.g., a new verb that recurred with food names might mean eat. Computational modelling (Resnik, 1995) and psycholinguistic studies (Trueswell, Tannenhaus, and Garnsey, 1994) suggest that such information can provide semantic-classificatory information about verbs. See Figure 5 for the example *call*.

Figure 5

Condition 3: Obervation and nominal context (Sequential semantic bootstrapping)

In this Condition, subjects received both the written list of nouns within sentence (as in Figure 5) and the matching videoclip for each such sentence (Figure 4). The procedure that we model here is one in which the learner has a stock of simple nouns, antecedently acquired (as in Experiment 1) by cross-situational observation, can recognize them in the speech stream being used with a novel verb, and checks this partial linguistic-contextual information against the ongoing scenes.

Condition 4: Syntactic Frames

For the subjects in this Condition, again there were no videotaped contexts. Instead, the mothers' sentences for the various mystery verbs appeared on a written list in partial nonsense form: All the nouns and verbs were converted to nonsense but the structural information (word order, and closed class words and morphemes) were left intact. The nonsense verb was capitalized. This Condition models a situation in which a learning device extracts the syntactic frame privileges of novel verbs, e.g., that see but not go occurs with tensed sentential complements; compare *I see that the cheese is in the refrigerator* with the awkward and non-occurring *I go that the cheese is in the refrigerator.* Figure 6 shows the stimuli for the example call.

Figure 6 -----

Condition 5: Syntax and selection

This Condition combined the noun (Figure 5) and structural (Figure 6) information by presenting the nouns of Condition 4 in their real rather than nonsense shapes, and in their original locations in the mothers' utterance, see Fig. 7 for the example *call*. Here again there was no video. Notice that this manipulation converted the co-occcurrence information of Conditions 2 and 3 to selectional information: The subject knew where in the sentence structure each meaningful noun occurred, and thus had significant clues to its thematic role.

Figure 7

Condition 6: Full information (Syntactic Bootstrapping)

In this Condition, subjects saw the video contexts (Figure 4) along with the sentences that accompanied each, with only the verb as nonsense (Figure 7). This Condition models the real input situation for learners, provided they have the internal wherewithal to represent it: repeated exposures to a word within syntactic structures, in company with the accompanying extralinguistic contexts.

Summary of the materials

On pain of being repetitious, we want to emphasize the stimulus properties of these six manipulations. For each Condition, the self-same extralinguistic situations and speech events sampled from several mothers provided the stimulus materials. Thus for call, the three mothers comprising the sample for this word, taken together, said it a total of seven times in our 3 hour sample. All seven videoclips of their doing so became the Condition 1 materials. The seven call-containing sentences that they then uttered yielded the nouns of Condition 2, the structures of Condition 4, and so forth. It is never the case that the video-clips come from one situation of a mother uttering call while the nouns or structures come from other situations in which she or another mother did so.

Subjects

One-hundred and twenty undergraduates participated in this experiment, 20 in each of the six Conditions. Subjects received credit in a psychology course or were paid.

Results and discussion

The results of all six Conditions are summarized in Tables 5 and 6. Table 5 shows the number and percentage correct for each Condition. As is obvious from inspection of the tabulated results, there are very large effects on identifiability as a function of the sources of information available. Table 6 reports a series of t-tests run on successive Conditions (1 - 6) to assess whether the outcome for each information source differed significantly from the successor source of information).[17] We now look in detail at these effects.

Tables 5 and 6

Learning from cross-situational observation (Condition 1)

As Column 1 of Table 5 shows, percent correct in the video condition was even lower (7.7%) than it had been in the roughly equivalent situation of Experiment 1 (15.3%). As we have earlier mentioned, the idea of using very short videoclips was so as to improve the time lock between utterance of the word and the relevant act (as assessed by the judges who created the tapes) and to lessen the potential confusion with temporally nearby events. But as the comparative results of the two versions show, this was counterproductive. Ss were aided far more by increasing the length of the context and thus gettiing bettter information about the gist of conversation, as in Experiment 1. This finding is consistent with Tomasello and Kruger (1992) who showed that very young children learn a new verb best if it is introduced when the event is impending rather than when it is already ongoing, and with Baldwin (1991) whose work shows that observers, even very young children, examine situations to discover something of the intents of speakers as part of their word-learning procedure. Thus the absolute level of performance in Experiment 1 for verb learning from context (15%) is probably a more reasonable estimate than the one achieved in the present manipulation.

On either estimate, however, the outcome was a very low success rate for verb identification if the observer's only recourse is to pair the passing scene with a single word (or beep). This experimental effect with adults is reminiscent of the inefficiency of verb learning early in the second year of life. "The world" that accompanies the utterance of a verb is too weak as a stand-alone information source for verb learning.

Another point of comparison between Experiment 1 and the present one is the intractability of particular verbs to learning by extralinguistic observation. Of the eight verbs in Experiment 1 that were identified by no Ss (<u>know, think, like, love, make, say, pop, see</u>), the six that reappear in the present condition (<u>know, think, like, say, make,</u> and <u>see</u>) were again uniformly misidentified. This supports the conclusion that concreteness or imageability rather than verbness *per se* is the operative variable.

<u>Noun co-occurrence information</u> (Condition 2)

A well-documented finding is that there are cross-linguistically stable prosodic cues to clause boundaries in the speech of mothers to infants (Fisher and Tokura, 1996). If so, then the very young learner, in advance of having control of the language-specific phrase structure, can establish a domain within which nouns (priorly learned via extralinguistic observation) stand in some constructional relationship to each other. Such a learner cannot directly know, however, which grammatical or thematic roles each of these nouns is playing in the sentence. This is because, e.g., the placement of the subject noun-phrase differs in the surface syntax of various languages.

The question raised in this Condition was whether these co-occurring nouns provide useable information as to the identity of the verb. Twenty new subjects saw lists of the nouns (in alphabetical order within sentence) that the mothers uttered with each of the 24 verbs, without being shown the videotaped events. Identifiability scores were higher than for Condition 1 (an 8.8% increase in success rate), though owing to variability this difference between the two Conditions did not reach significance ($p = .10$).

It is something of a surprise that noun co-occurrence information, taken alone, was as useful as it was (16.5% mean correct, see Table 6). After all, without knowing just where (structurally) the nouns actually occurred within sentences, in theory one cannot make very secure inferences about the verbs (carrots are eat<u>en</u>, but the rabbits are their eat<u>ers).</u> Moreover, some verbs seem to accept just about any nouns (you can <u>find</u> anything loseable and <u>throw</u> anything you can lift).[18] Finally, in mother-to-baby talk, nouns very often are names and pronouns, items licensed for just about any verb and therefore informative for none of them; see Figure 5.

Then why did Ss exhibit some measure of success under these information conditions? For one thing, a noun like *phone* can be a give-away to a small class of common verbs, such as *talk*, *listen,* and *call*. Also, different subcategories of nouns probabilistically perform different thematic roles, a factor whose influence can be observed even in rapid on-line parsing performance (Trueswell, 1996). As one major example, animate nouns are vastly more likely than inanimates to appear in subject position just because they are likely to be the causal agents in events. More generally, various "prominence" factors determine which noun will capture the subject position. [19]

The sheer number of nouns in the maternal sentence provides an additional clue to the verb meaning. This is because in the very short sentences used to infants, the number of nouns is a soft indicator of the number of arguments (Fisher et al, 1994, Fisher, 1996; Naigles, 1990; Naigles, Gleitman and Gleitman, 1992; Naigles, Fowler, and Helm, 1992). Thus *gorp* in *John is gorping* is more likely to mean 'sneeze' than 'kick.' And 'kick' is a better guess than 'sneeze' for either *John is gorping the snaggle* or *The snaggle is gorping John* even if, because of lack of syntactic knowledge, one cannot tell one of these last two from the other.

Observation and noun context (Condition 3)

This Condition modeled a learner who can coordinate the two sources of evidence we have so far discussed, inspecting the world to extract salient conjectures about relevant events, and using the known nouns to narrow the choice among them. This hypothetical procedure is a version of the "semantic bootstrappping" schemes of Grimshaw (1981) and Pinker (1984), only it is sequential (i.e., the search for the verb interpretation takes place in the context of antecedently acquired nouns whose occurrence the learner can register). Accordingly, in this manipulation subjects received the noun lists (as in Figure 5) and could take these into account while inspecting the scenes (as in Figure 4). Armed with these dual information sources, for the first time subjects achieved a respectable level of identification of new verbs (29%, a statistically significant improvement from performance where only scenes or only nouns were available, $p > .01$).

Syntactic frame Information (Condition 4)

In this Condition, we assessed the informativeness of distributional information concerning the structures in which verbs appear in the exposure language, their subcategorization privileges. Despite the manifest oddity of seeing merely a written list of these naked nonsense frames (but see also Lewis Carroll, 1865) without extralinguistic context (no video), subjects in this Condition identified 51.7% of the verbs. This represents a dramatic improvement ($p < .01$) even over the dual-cue Condition 4. All the information from noun knowledge and from observation of the passing scene was withheld from the subjects, and yet their performance level leapt up. Why?

Syntactic information can cue to verb meaning just because the structural privileges of a verb (the number, type, and positioning of its associated phrases) derive , quirks and provisos aside, from its argument-taking properties. The number of argument positions lines up with the number of participants implied by the logic of the verb. Thus a verb that describes a self-caused act of the musculature (e.g., Joe *snoring*) is liable to surface intransitively, a physical effect of one entity on another (Joe *throwing* a ball) is likely to be labelled by a transitive verb, and an act of transfer of an entity between two places or persons is likely to be ditransitive (Joe *giving* a ball to Bill). The type of complement is also derivative of aspects of the verb's meaning. Thus a verb describing a relation between an actor and a proposition is likely to take clause-like complements (Joe *believing that Bill is sad*). Because verb meanings are compositional at least at the level of these argument-taking properties (Grimshaw, 1990), the matrix of verb-to-structure privileges has the effect of providing a coarse semantic partitioning of the verb set. For example, because one can forget *things*, this verb licenses a noun-phrase complement; and because one can also forget *events*, it also licenses clausal complements. A vast linguistic literature documents these syntax-semantics relations (see, e.g., Gruber, 1967, Fillmore 1968, and McCawley, 1968 for seminal discussions, Croft, 1991, Goldberg, 1995, and Levin, 1993 for recent treatments; for experimental documentation, Fisher, Gleitman, and Gleitman, 1991; Lidz, Gleitman, and Gleitman,1998; and for cross-linguistic evidence concerning caretaker speech, Lederer, Gleitman, and Gleitman, 1995; Geyer, 1998; and Li, 1994; for learning effects in young children, Bloom, 1994; Brown, 1957; Fisher et al., 1994; Mintz and Gleitman, 1998; Naigles, 1990; 1996; Naigles, Fowler, and Helm, 1995; Waxman, 1994). In this Condition, subjects were able to use syntax and morphology to make inferences about the verb meanings even though they were artificially disbarred from observing the contexts of use and the cooccurring nouns.

Finally, notice that the types of verb that are easily identified from situational and syntactic evidence differ. The 12 verbs *never identified* in Condition 1 (extralinguistic observation) are the 12 that are *more easily identified* in the present syntactic condition: These 12 "more abstract" verbs were identified via the syntactic clues 65.4% of the time whereas the "more observable" ones were identified by syntax only 37.9% of the time ($t = -2.31$, $df = 22$, $p < .04$). Experimental variability aside, the real difference here is between verbs whose content is mental (see, look, want, know, like, think) vs those that encode physical, observable action: These mental verbs are correctly identified 90% of the time via syntactic evidence while all the other verbs are collectively identified by this evidence only 40% of the time ($t = 4.59$, $df = 22$, $p < .0001$).

All this stands to reason. It would have to be the case that for relations that are not available for perceptual inspection -- such as the invisible, odorless thoughts and desires in other minds -- there simply *must* be reliable alternatives to extracting them from the world in view. Otherwise, how could they be learned at all?

Syntax and selection (Condition 5)

In this Condition, we modelled a device that has access to both syntactic (Condition 4) and noun (Condition 2) information, i.e., to full sentences in which only the verb is unidentified. Notice that now the co-occurence information (the association of food nouns with ingestive verbs, for example) has been converted to selectional information; that is, knowledge of which noun occurred in which structural position. This ought to be, and is, a material improvement in the quality of information. Now the subjects identified the targets 75% of the time, a 24% improvement over performance in Condition 4 ($p < .01$). For after all, a child who hears the food words in sentences like "That *candy* will ruin your appetite" is ill-served if she takes this mere proximity as evidence that *ruin* means '*eat.*' *Candy* in object position, though still no infallible cue to a unique verb identity (e.g., "Do you *want* some candy?"), is a much more secure piece of evidence.

The real input to verb learning: (Condition 6)

In this final manipulation, we made available to our subjects several sources of information that we believe are available to normally circumstanced 2-year learners: adult utterances in supportive extralinguistic contexts. In this informational Paradise, the subjects approached perfect identification at 90.4%[20]. We can begin to understand the quick and relatively errorless child vocabulary feats by assuming that, like the adults in this experiment, they make use of multiple, mutually constraining, evidentiary sources.

Part III: Discussion

> If we will observe how children learn languages, we shall find that...people ordinarily show them the thing of which they would have them have the idea; and then repeat to

> them the name that stands for it, as ‘white’, ‘sweet’, ‘milk’, ‘sugar’, ‘cat’, ‘dog’.
>
> John Locke, 1690; Book 3.IX.9)

The experiments we have reported were attempts to discover whether word learning is as straightforward as John Locke and many of his descendants in developmental psycholinguistics have supposed. Can the meanings of words really be derived by a process that has access solely to the extralinguistic contexts for their use? To find out, we radically reduced the task to one in which adults merely have to *identify* words that they have known since their earliest childhood, based on such word-to-world patterning. Perhaps the primary finding of this work was that the environment does not seem to be so simply and generally informative after all. This is just because the same observations make available a variety of construals for any single word.

The second major finding was that the extralinguistic contexts were far more informative for some kinds of words than for others. On first assessment, it appeared that this distinction was between the lexical classes noun and verb. But the most greatest identification difficulties were for a subset of the verbs, those that refer to mental states and acts, such as *think* and *see.* These were never identified when all the subjects had to go on was the visual-contextual information. On reflection, this makes perfectly good sense. For observation to help link a word to an event or state of affairs, this aspect of the world has to be observable. Indeed, further analysis showed that the main predictor of identifiability was not lexical class as such. Instead, it was a function of the words’ concreteness and imagibility.

The third result concerned the power of syntactic cues for identifying the verbs, especially for just those mental-content verbs for which the extralinguistic video information had been least informative. Finally, and most generally, the outcomes were that a variety of cues -- environmental and linguistic -- contributed probabilistically to identification of the verbs. Despite the many artificialities of our tasks, and the limitation of our subjects to a few exemplars of what it is they were to learn, when provided simultaneously with several kinds of linguistic and situational evidence their “learning” of even the most abstract of the maternal verbs was at ceiling.

We now ask in further detail how the Human Simulation paradigm might be useful in thinking about the child’s acquisition of a first lexicon. After all, our results were for adult subjects trying to identify words that they actually had aquired 20 or so years before. How do these bear on the ways in which young children manage to map sound categories onto meanings, such that they can use words to refer to objects and events in the world?

Conceptual and linguistic growth

In the Introduction to this paper, we approached the problem of understanding vocabulary acquisition by pointing to the categorial limitations on words young children first acquire. Adults utter words from all the lexical classes , but the infants' earliest vocabulary overwhelmingly consists of nouns. We now reconsider two interpretations of this oddity. The first invokes a postulated difference between the conceptual structures of adults and young children. Pinker, 1984, called this *the discontinuity hypothesis* for it holds that children are organizing the world in terms of mental representations fundamentally different from those that characterize adults. An instance of this perspective is Gentner (1978) which describes the noun advantage as arising from a conceptual-complexity distinction in the typical ways these words refer to the world, nouns typically describing object-reference concepts, and verbs labelling the relations between the objects. According to this view, then, learning words is not only a matter of mapping between concepts and sound categories; it is a matter of concept learning as well. To the extent that this is correct, there would be no reason to expect adult word-identification performance in Experiment 1 of the present series to reproduce the noun-first characteristic of child word learning, for these adults have manifestly attained the conceptual wherewithal to understand both object and relational terms.

The contrasting approach is called *the continuity hypothesis.* In regard to the word-learning question, it assumes that children are conceptually equipped to entertain the concepts encoded by most of the words that their caregivers say to them, nouns and verbs alike. The children's task is (merely!) to discover which sound patterns in the language map onto which of these meanings (for this view in pristine form, Fodor, 1981). Within the terms of this latter hypothesis, it is still possible to account for why children are at first heavily restricted to noun learning. This is by positing that different kinds of words require different kinds of information to acquire.

It has been hard to adjudicate between these two positions *in vivo*, so to speak, for the infant learners may be undergoing both linguistic and cognitive growth at the same time. The rationale for studying adults was to remove potential conceptual factors from the equation so as to understand the contribution of linguistic-informational factors to vocabulary growth. The findings clearly favor an information-based account, as we now discuss in further detail.

Schematizing the word-learning machinery

The results of Experiment 3 suggest that efficient vocabulary acquisition requires the recruitment of several kinds of linguistic and extralinguistic information. The assignment of interpretation is the one that best satisfies the constraints contributed by each of these input sources. However, in generalizing from the adult behavior to the child learning problem we run the risk of begging our own question, for how -- if not by learning the word meanings first -- do children come into possession of the linguistic structures now claimed to be part of the critical *input* to word learning? (See Pinker, 1994, for this kind of protest) . As we will describe, the circle is avoided because the novice breaks into the system by a first-pass, asyntactic, procedure whose output is a concrete nominal vocabulary. This primary knowledge underpins the construction of clause level syntax -- which in turn enables further vocabulary growth.

Concrete nouns as the scaffold for language learning

Necessarily, vocabulary acquisition begins as Locke supposed: with the child's attempt to discover contingencies between sound categories and recurrent aspects of the world. The output of this procedure is a small stock of routinized social expressions ("bye-bye") and common nouns. As we showed, there is a very prosaic reason why this should be so, not only for the youngest language learners (Table 1) but also for adults whose information is artificially restricted to this procedure: Not every lexical item encodes a category that can be physically (no less "saliently") instantiated in the world, and "observed." The very frequent nouns used by mothers of young children are as a group more straightforwardly observable than their frequent verbs. In Experiment 2 of the present series, allmost all the verbs were rated as lower in imageability than any of the nouns even though there were differences in imageability within each of these lexical classes (Table 3). So strong was the imageability factor that it completely swamped the lexical-class factor. That is, nouns were easier to identify by word-to-world pairing just because "being a noun" and "being highly imageable" were virtually the same thing for these materials.

This is a particularly useful finding, because there would be many reasons to be dismayed if there were a *principled* learning distinction for noun versus verb learning. The first of these is that we want to be able to explain why in the earliest child vocabularies verbs, while rare, are not altogether absent. Relatedly, the same explanation that accounts for why a stock of nouns is learned first should account as well for why the first verbs learned (both by infants and by our adult subjects) are concrete ones like *throw* and not abstract ones like *want*. For that matter, by much the same explanatory apparatus we want to account for why even Fido seems able to acquire certain verbs like *roll over* (as well as nouns like *bisquit)* despite this creature's indifference to English syntax. Finally, we need a theory that can accomodate to the finding that the noun-dominance property of early speech and comprehension, though enormously robust, is sensitive to properties of input that vary across language and across culture (see note 2). The obvious solution, instantiated in the adult performances reported here, is that learning solely by examining extralinguistic context supports discovery of just those words whose instances are yielded up straightforwardly by perception. In practice, this will heavily favor words that label object-reference concepts, and these are generally nouns.

Despite the manifest limitations of word-to-world pairing, both for novices and sophisticated college sophomores it must be the rock-bottom foundation of vocabulary acquisition. Later developments do not and could not materially diminish the role that extralinguistic observation must play. *But only for some core set of object-reference terms does this procedure have to operate in the relative absence of other supporting cues to word interpretation.*

Co-occurrence information and argument structure

Table 5 shows the full results of Experiment 3, namely the dramatically differing success rates for verb identification as we varied the subjects' access to information types. Conditions 2 (noun co-ocurrence) and 3 (co-occurrence + scenes) showed that these "learners," when provided with information about which nouns cooccurred with the mystery verbs, could increment their learning efficiency in several ways. Most obviously, they could draw verb-semantic inferences by examining the plausibility of certain verb meanings given the nouns. For example, when balls are mentioned, *throwing* is a salient relation (for this effect with 15-month old one-word speakers, Shipley, Smith, and Gleitman, 1969). Moreover, the number of nouns probabilistically cues the number of arguments in the predicate. This roughly distinguishes between binary relations (such as *push,* e.g., *"The kangeroo pushed the monkey"*) and unary relations (such as *fall*, e.g., *The monkey fell*) in the many common situations where observation makes both interpretations available (for this effect with one- and two- word speakers, Naigles, 1990; for a discussion of the power of even unlabelled phrase-structure trees for grammatical inference procedures, Joshi and Levy, 1982; Fisher et al, 1994). When the referential (video) information was made available along with the nouns (Condition 3), the subjects identified a respectable 29% of the mystery verbs. That is, the extralinguistic contexts become more useful for verb learning when their interpretation is reined in by inspecting their deployment with respect to the priorly acquired nouns.

Noun-cooccurrence information is useful beyond its role in identification. The partial information from noun-verb coocccurence can help in the process of acquiring the phrase structure itself. This is because, as we have discussed earlier, inherent inequalities in the prominence of the nouns in the clause (in animacy, prototypicality, and causal role) provide a crucial clue to discovery of the surface position of the subject of the sentence in the exposure language (Dowty, 1991; Ertel, 1977; Fisher et. al., 1994; Fisher, 1996; Grimshaw, 1990; Osgood and Bock, 1977; Talmy, 1975). Once the position of the sentence subject is known, there is little further clause-level syntax for the learner to acquire.

Structural supports for verb learning

Accuracy of the verb-identification process improved very significantly when subjects were provided with syntactic context, even when (as in Condition 4) extralinguistic and noun-cooccurrence information were withheld; under this presentation condition, subjects identified over half the verbs. This effect is really not surprising. Verbs are argument-taking predicates. Information as to this argument structure is displayed across the sentence in terms of its phrase structure and associated morphology (We allude here to generalizations which are, under slightly varying formalizations, collectively referred to as the Projection Principle and the Theta Criterion, Chomsky, 1981).

Two kinds of evidence in the literature bolster the position that syntactic information aids child learners in much the same way as it did these adults. First, we have shown elsewhere that mothers' choices of verb-subcategorization environments in speech to their infants partitions the verb class semantically in the required way: Overlap in subcategorization privileges predicts semantic relatedness (Lederer et al., 1995) and, for major cases, does so in the same way for maternal speech in languages as disparate as English, Mandarin Chinese (Li, 1994), and Hebrew (Geyer, 1998). Second, there is now indisputable evidence in the literature that toddlers actively recruit such evidence and use it to constrain their search for the meanings of novel words (Bloom, 1994; Fisher et al., 1994; Fisher, 1996; Hirsh-Pasek and Golinkoff, 1991; Naigles, 1990, 1996; Naigles and Kako, 1993).

Of course, syntactic evidence taken alone cannot divulge the meaning of individual verbs. There is no "hot syntax" or "cold syntax" that will differentiate *freeze* from *burn*. Only the verb's argument-taking properties can be revealed by its syntactic privileges of occurrence, such as the syntactic subcategorization information made available in Condition 4. But the power of this information source was such that subjects now identified 52% of the mystery words -- a far stronger performance than the 29% success rate achieved in Condition 4 which supplied the videotaped scenes and their accompanying nouns.

A subresult of the Condition 4 findings is particularly relevant in understanding the evolving machinery of vocabulary acquisition. In the presence of syntactic information, and syntactic information *only*, subjects were significantly better at identifying abstract verbs like *think* than concrete verbs like *go*. This finding is nicely symmetrical wit h the findings of Experiment 1 and 2: Where observation is the only source of evidence, the most concrete words (in practice, the nouns) have the advantage; where syntax is the only source of evidence, the least concrete words (the mental-content verbs) are the easiest. Evidently, where observation provides the least information for learners, the syntactic (subcategorization) distinctions made by the language are the most precise. After all, if thinking and knowing cannot be directly observed, some other means have to be supplied, else words labelling these distinctions could not be acquired.

In the real case of learning children as simulated in Conditions 5 (syntax + nouns, i.e., the maternal sentences without video) and 6 (syntax + nouns + video), the constraints on interpretation made available by attention to verb-specific syntactic information were offered conjointly with constraints from other sources. Under these conditions, the identification of the verbs was a snap, with subjects identifying 90% of the verbs in Condition 6, virtually ceiling performance. The developmental literature documents that analogous steps have been carried out by normally developing children well before the second birthday. Crucially, seventeen-month old children have language-specific knowledge of the positioning of the subject noun-phrase and its semantic value; for instance, they will appropriately fixate on different pictures depending on whether they hear *Big Bird is tickling Cookie Monster* or *Cookie Monster is tickling Big Bird* (Golinkoff, 1975; Hirsh-Pasek and Golinkoff, 1991). Once the parse-tree representations become available, learning of verbs and, indeed, words from every lexical class, is overdetermined. The experimental upshot is that subjects offered such information have no trouble identifying the mystery words. The upshot in the real world is that by 5 years of age, children have vocabularies of 10 to 15,000 words which sample the lexical classes more or less as predicted by their frequency ranges in adult-to-child speech.

The Human Simulation procedure thus bolsters and solidifies the interpretation of a variety of fragmentary results from observational and experimental studies with young children. Babies who show no evidence of language-specific syntactic knowledge comprehend and produce mainly concrete nouns, while verbs begin to make their appearance at the first evidence of rudimentary syntactic knowledge (Caselli et al., 1995; Gentner, 1978; Shipley et al., 1969). At the onset of the two-word stage in language knowledge, infants add to their arsenal of learning tricks the various distributional and syntactic analyses necessary to constrain the meanings of all but the simplest first word. As we showed in Experiment 3, the less "concrete" the items to be acquired (i.e., the more their semantic content implicates the unob-servable mental states and acts of their users) the greater is the reliance of the learner on nonre-ferential, language-internal, cues to the meanings.

IV. FINAL THOUGHTS

We conceived of the present experiments as support for an incremental constraint-satisfaction view of how children acquire their native tongue (for related discussion, see also Kelly and Martin, 1994; Saffran et al, 1996; Seidenberg and MacDonald, in press). We believe that this work bolsters the considerable accumulating evidence that lexical and syntactic knowledge, far from developing as separate components of an acquisition procedure, interact with each other and with the observed world in a complex mutually supportive series of bootstrapping operations whose outcome is a lexicalized grammar. Adults were the population of choice because they allowed us to explore information-based properties of this learning procedure apart from conceptual growth factors.

In no way, of course, do the findings rule out the idea that conceptual factors could *also* be implicated in the course that word and syntax learning take during the first three years of life. Indeed Gentner's (1978; 1981) experimental analysis of many memorial, representational, and processing differences between nouns and verbs lends a good deal of credence to the idea that concept typology is playing a role in the character of early vocabularies. Still, we have provided an existence proof that even when these issues are excluded by experimental artifice, nouns are acquired more efficiently than verbs , and action verbs more efficiently than mental verbs, if the information base is limited to extralinguistic observation. Once the full system of cues is in place, words like *know* are no harder to acquire than words like *put* (perhaps easier!).

In sum, we began by noting that the word-learning achievements of infants are limited to the homeliest terms labelling the ordinary objects and creatures in their ambient world. One explanation is that infants are conceptually primitive. But our findings suggested that when deprived of the use of the distributional and syntactic properties of their native tongue, adults too are slow, errorful, and restricted word learners. The key to efficient word learning is to satisfy the several constraints available from sophisticated extralinguistic *and* linguistic representations of conversational exchange. Though each such cue is fallible and incomplete in the information provided, together they lead to convergence on the interpretations of the lexical items. It could well be, then, that infant learners are conceptual creatures much like us, able to contemplate not only the dogs and the spoons and the balls, but also their own and others' thoughts, beliefs, and desires. Yet, owing to the fact that one-year olds do not know English (or Greek, or Malukakan, or any other other special instantiation of the human language program), their earliest labelling can hardly rise beyond the concrete terms.

BIBLIOGRAPHY

Au, T.K., Dapretto, M. & Song, Y.K. (1994). Input versus constraints: Early word acquisition in Korean and English. *Journal of Memory and Language*, **33**, 567-582.

Baldwin, D.A. (1991). Infant contributions to the achievement of joint reference. *Child Development*, **62**, 875-890.

Bates, E., Dale, P.S., & Thal, D. (1995). Individual differences and their implications for theories of language development. In P. Fletcher & B. MacWhinney (Eds.), *Handbook of Child language* (pp. 96-151). Oxford: Basil Blackwell.

Bloom, P. (1994). Possible names: The role of syntax-semantics mappings in the acquisition of nominals. *Lingua*, **92**, 297-329.

Brent, M.R. (1994), Surfrace cues and robust inference as a basis for the early acquisition of subcategorization frames, in L. Gleitman and B. Landau (eds.) pp. 433-470, *The acquisition of the lexicon*, Cambridge, MA: MIT Press.

Brent, M.R. and Cartwright, T.A. (1996) Distributional regularity and phhonotactic constraints are useful for segmentation, *Cognition* **61 1:2,** 93-125.

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

Brown, R. (1957). Linguistic determinism and the part of speech. *Journal of Abnormal and Social Psychology*, **55**, 1-5.

Brown, R. (1973). *A first language*. Cambridge, MA: Harvard University Press.

Bruner, (1974/1975). From communication to language: A psychological perspective. *Cognition*, **3**, 255-287.

Carey, S. (1978). The child as word learner. In M. Halle, J. Bresnan, & A. Miller (Eds.), *Linguistic theory and psychological reality* (pp. 264 - 293). Cambridge, MA: MIT Press.

Carey, S. (1994) Does learning a language require a child to reconceptualize the world? In L. Gleitman and B. Landau (eds.), pp. 143-168, *Acquisition of the lexicon*, Cambridge, MA: MIT Press.

Carroll, L. (1865). *Alice's adventures in wonderland*. New York: Bantam Books.

Cartwright, T.A., & Brent, M.R. (1997). Syntactic categorization in early language acquisition: Formalizing the role of distributional analysis. *Cognition*, **63**, 121-170.

Caselli, M.C., Bates, E., Casadio, P., Fenson, J., Fenson, L., Sanderl, L., & Weir, J. (1995). A cross-linguistic study of early lexical development. *Cognitive Development*, **10**, 159-199.

Choi, S. & Gopnik, A. (1995). Early acquisition of verbs in Korean: A cross-linguistic study. *Journal of Child Language*, **22**, 497-529.

Chomsky, N. (1981). *Lectures on the theory of government and binding*. Foris: Dordrecht.

Croft, W.A. (1991), *Syntactic categories and grammatical relations*, Chicago, IL: Univ. Of Chicago Press.

Dowty, D.R. (1991). Thematic proto-roles and argument selection, *Language*, **67**, 547-619.

Elman, J. (1993). Learning and development in neural networks: The importance of starting small, *Cognition* **48:1,** 71-99.

Ertel, S. (1977). Where do the subjects of sentences come from. In S. Rosenberg (Ed.), *Sentence production: Developments in research and theory*. Hillsdale, NJ: Erlbaum.

Fillmore, C. (1968). The case for case. In E. Bach & R.T. Harms (Eds.), *Universals of linguistic theory*, NY: Holt, Reinhart, and Winston.

Fisher, C. (1996). Structural limits on verb mapping: The role of analogy in chil dren's interpretations of sentences. *Cognitive Psychology*, **31**, 41-81.

Fisher, C., Gleitman, H., & Gleitman, L.R. (1991). On the semantic content of subcategorization frames. *Cognitive Psychology,* **23**, 331-392.

Fisher, C., Hall, G., Rakowitz, S., & Gleitman, L. (1994). When it is better to receive than to give: Syntactic and conceptual constraints on vocabulary growth. *Lingua*, **92**, 333-375.

Fisher, C. & Tokura, H. (1996). Prosody in speech to infants: Direct and indirect acoustic cues to syntactic structure. In J.L. Morgan & K. Demuth (Eds.), *Signal to syntax: Bootstrapping from speech to grammar in early acquisition* (pp. 343-363). Mahwah, NJ: Lawrence Erlbaum Associates.

Fodor, J. (1981) The present status of the innate controversy, in J. Fodor , *Representations,* 257-316, Cambridge MA: MIT Press.

Gati, I. & Tversky, A. (1984). Weighting common and distinctive features in perceptual and conceptual judgments. *Cognitive Psychology*, **16**, 341-370.

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

Gentner, D. (1981). Some interesting differences between verbs and nouns. *Cognition and Brain Theory,* **4**, 161-178.

Gentner, D. (1982). Why nouns are learned before verbs: Linguistic relativity versus natural partitioning. In S.A. Kuczaj, II (Ed.), *Language development: Vol. 2. Language, thought, and culture* (pp 31-53). Hillsdale, NJ: Erlbaum.

Gentner, D., & Boroditsky, L. (in press). Individuation, relativity, and early word learning. In M. Bowerman & S. Levinson (Eds.), Language acquisition and conceptual development. England: Cambridge University Press.

Gentner, D. & France, I.M. (1988). The verb mutability effect: Studies of the combinatorial semantics of nouns and verbs. In S. Small, G. Cottrell, & M. Tannenhaus (Eds.), *Lexical ambiguity resolution: Perspectives from psycholinguistics, neuropsychology, and artificial intelligence* (pp. 343-382). San Mateo, CA: Kauffman.

Gentner, D., & Ratterman, M. (1991). Language and the career of similarity. In S.A. Gelman & J.P. Byrnes (Eds.), Perspectives on language and thought: Interrelations in development. Cambridge: Cambridge University Press.

Geyer, H. (1998). Subcategorization as a predictor of verb meaning: Evidence from Hebrew. Unpublished ms., University of Pa..

Gleitman, L. (1990). The structural sources of word meaning. *Language Acquisition*, **1**, 3-55.

Gleitman, L. and Gleitman, H. (1997) What is a language made out of?, *Lingua* **100,** 29-55.

Gleitman, L., Gleitman, L., Landau, B., and Wanner, E. (1988), Where learning begins: initial representations for language learning, pp. 150-193, in F.J. Newmeyer (ed.), *Linguistics: The Cambridge survey*, Vol. III, Cambridge: Cambridge Univ. Press.

Gleitman, L.R., Gleitman, H., Miller, C., & Ostrin, R. (1996). Similar, and similar concepts. *Cognition*, **58,** 321-376.

Gleitman, L., & Wanner, E. (1982). Language acquisition: The state of the art. In E. Wanner & L. Gleitman (Eds.), *Language acquisition: The state of the art* (pp. 1-50). New York: Cambridge University Press.

Goldberg, A.. (1995) *Constructions: A construction grammar approach to argument structure,* Chicago: Univ. Of Chicago Press.

Goldin-Meadow, S., Seligman, M.E.P., & Gelman, R. (1976). Language in the two-year-old. *Cognition*, **4**, 189-202.

Goldowsky, B.N. & Newport, E.L. (1996). Modeling the effects of processing limitations on the acquisition of morphology: The Less is More Hypothesis. In j. Mead (ed.), *Proceedings of the 11th West Coast Conference on Formal Linguistics,* Stanford CA.

Golinkoff, R. (1975) Semantic development in infants: The concepts of agent and recipient, *Merrill-Palmer Quarterly*, **21**, 181-193.

Golinkoff, R.M., Jacquet, R.C., Hirsh-Pasek, K., & Nandakumar, R. (1996). Lexical principles may underlie the learning of verbs. *Child Development*, **67**, 3101-3119.

Grimshaw, J. (1981). Form, function, and the language acquisition device. In C.L. Baker & J.J. Baker (Eds.), *The logical problem of language acquisition*. Cambridge, MA: MIT Press.

Grimshaw, J. (1990). *Argument structure*. Cambridge, MA: MIT Press.

Gruber, J.S. (1967). Look and see, *Language*, **43**, 937-947.

Hirsh-Pasek, K. & Golinkoff, R. (1991). Language comprehension: A new look at some old themes. In N. Krasnegor, D. Rumbaugh, M. Studdert-Kennedy, & R. Schiefelbusch (Eds.), Biological and behavioral aspects of language acquisition (pp. 301-320). Hillsdale, NJ: Lawrence Erlbaum Associates.

Hirsh-Pasek, K., Golinkoff, R., Gleitman, H., Gleitman, L., and Naigles, L. (1988, October) *Syntactic bootstrapping: Evidence from comprehension*, Paper presented at the Boston Conference on Language Development, Boston, MA.

Hume, D. (1738). *An inquiry concerning human understanding*. Indianapolis, IN: Bobbs-Merrill, 1955.

Huttenlocher, J., Smiley, P., & Charney, R. (1983). The emergence of action categories in the child: Evidence from verb meanings. *Psychological Review*, **90**, 72-93.

Huttenlocher, J., Smiley, P., & Ratner, H. (1983). What do word meanings reveal about conceptual development? In T.R. Wannenmacher & W. Seiler (Eds.), *The development of word meanings and concepts* (pp. 210-233). Berlin: Springer-Verlag.

Joshi, A.K. & Levy, L.S. (1982). Phrase structure trees bear more fruit than you would have thought. *American Journal of Computational Linguistics*, **8**, 1-11.

Juszyck, P. (1997) The discovery of spoken language. Cambridge MA: MIT Press.

Kelly, M. and Martin, S. (1994) Domain-general abilities applied to domain-specific tasks: Sensitivity to probabilities in perception, cognition, and language, in L. Gleitman and B. Landau (eds.), *The acquisition of the lexicon*, 105-142, Cambridge, MA: MIT Press.

Landau, B. (1994). Object shape, object name, and object kind: Representation and development. In D.L. Medin (Ed.), *The psychology of learning and motivation: Advances in research and theory, Vol. 31*. San Diego, CA: Academic Press, Inc.

Landau, B. & Gleitman, L.R. (1985). *Language and experience: Evidence from the blind child*. Cambridge, MA: Harvard University Press.

Lederer, A., Gleitman, H., & Gleitman, L. (1995). Verbs of a feather flock together: Semantic information in the structure of maternal speech. In M. Tomasello & W. Merriman (Eds.), *Beyond names for things* (pp. 277-297). Hillsdale, NJ: Erlbaum.

Lennenberg, E.H. (1967). *Biological foundations of language*. New York: Wiley.

Levin, B. (1993). *English verb classes and alternations*. Chicago: University of Chicago Press.

Levine, S., & Carey, S. (1982). Up front: The acquisition of a concept and a word. Journal of *Child Language*, **9**, 645-658.

Li, P. (1994) Maternal verb usage in Mandarin Chinese, Unpublished ms., Univ. of Pa.

Lidz, J., Gleitman, H. & Gleitman, L. (1998) Reference to come.

Locke, J. (1690). *An essay concerning human understanding* (pp. 259-298). Cleveland: Meridian Books, 1964.

MacNamara, J. (1972). Cognitive basis of language learning in infants. *Psychological Review*, **79**, 1-13.

Mandler, J. (1992) How to build a baby II: conceptual primitives. *Psychological Review.* **99**, 587-604.

Marler, P. (1970).A comparative approach to vocal learning: Song development in white crowned sparrows. *Journal of Comparative and Physiological Psychology*, monograph 7, 1-25.

McCawley, J.D. (1968) The role of semantics in a grammar, in E. Bach and R. Harms (eds.), 124-169, *Universals in linguistic theory*, NY: Holt, Reinhart, and Winston.

Merriman, W., & Tomasello, M. (1995). Introduction: Verbs are words too. In M. Tomasello & W.E. Merriman (Eds.), *Beyond Names for Things* (pp. 1-18). Hillsdale, NJ: Lawrence Erlbaum Associates.

Miller, C.A. (1998). It takes two to tango: understanding and acquiring symmetrical verbs. *Journal of Psycholinguistic Research*, **27**, 385-411.

Mintz, T., & Gleitman, L. (1998) Incremental language learning: two and three year olds' acquisition of adjectives. *Proceedings of the 20th Annual Conference of the Cognitive Science Society*, Hillsdale, NJ: Erlbaum.

Mintz, T., Newport, E. and Bever, T. (1995). Distributional regularities of form class in speech to young children. In J. Beckman (ed.), *proceedings of the 25th Meeting of the Northeastern Psychological Society*, Hillsdale, NJ: Erlbaum.

Morgan, J.L., Meier, R.P., & Newport, E.L. (1989). Facilitating the acquisition of syntax with cross-sentential cues to phrase structure. *Journal of Memory and Language*, **28**, 360-374.

Naigles, L. (1990). Children use syntax to learn verb meanings. *Journal of Child Language*, **17**, 357-374.

Naigles, L. (1996). The use of multiple frames in verb learning via syntactic bootstrapping. *Cognition*, **58**, 221-251.

Naigles, L., Fowler, A., & Helm, A. (1992) Developmental shifts in the construction of verb meanings. *Cognitive Development*, **7**, 403-428.

Naigles, L. & Kako, E. (1993). First contact in verb acquisition: Defining a role for syntax. *Child Development*, **64**, 1665-1687.

Nelson, K. (1981). Individual differences in language development: Implications for development and language. *Developmental Psychology*, **17**, 170-187.

Newport, E. (19 90). Maturational constraints on language learning. *Cognitive Science* **14**, 11-28.

Newport, E., Gleitman, H., & Gleitman, L. (1977). Mother, I'd rather do it myself: Some effects and noneffects of maternal speech style. In C.E. Snow and C. A. Ferguson (Eds.), *Talking to children: Language input and acquisition* (pp.109-50). Cambridge: Cambridge University Press.

Osgood, C.E. & Bock, J.K. (1977). Salience and sentencing: Some production principles. In S. Rosenberg (Ed.), *Sentence production: Developments in research and theory* (pp. 89-145). Hillsdale, NJ: Erlbaum.

Pinker, S. (1984). *Language Learnability and Language Development*. Cambridge, MA: Harvard University Press.

Pinker, S. (1994) How could a child use verb syntax to learn verb semantics? In L. Gleitman and B. Landau (eds.), pp 377-410, *The acquisition of the lexicon*, Cambridge MA: MIT Press.

Resnik, P. (1995) Selectional constraints: an information-theoretic model and its computational realization;, *Cognition* **61**, 1-2, 127-159.

Saffran, J.R., Newport, E.L., & Aslin, R.N. (1996). Word segmentation: The role of distributional cues. *Journal of Memory and Language*, **35**, 606-621.

Seidenberg, M. & McClelland, J. (1989). A distributed, developmental model of word recognition and naming. *Psychological Review* **96**, 523-568.

Seidenberg, M. And MacDonald, M.C. (in press) A probabilistic constraints-based approach to language acquisition and processing, *Cognitive Science*, special issue, *Connectionism,* M. Christiansen, N. Chater, and M. Seidenberg (eds.).

Shipley, E.F., Smith, C., & Gleitman, L.R. (1969). A study in the acquisition of language: free responses to commands. *Language*, **45**, 322-42.

Smiley. P. and Huttenlocher, JE. (1995) Conceptual development and the child's early words for events, objects, and persons, in M. Tomasello and W. Merriman (eds.), *Beyond names for things*, 21-62, Hillsdale, NJ: Erlbaum.

Soja, N.N., Carey, S. & Spelke, E.S. (1991). Ontological categories guide young children's inductions of word meaning: object terms and substance terms. *Cognition*, **38**, 179-211.

Spelke, E.S. (1985) Perception of unity, persistence, and identity: Thoughts on infants' conceptions of objects. In J. Mehler, R. Fox (eds.), *Neonate cognition,* Hillsdale, NJ: Erlbaum.

Talmy, L. (1978) The relation of grammar to cognition. In D. Waltz (ed) *Proceedings of TINLAP-2* Champaign: Coordinated Science Laboratory, Univ. of Ill.

Talmy, L. (1978). Lexicalization patterns: Semantic structure in lexical forms. In T. Shopen (Ed.), Language, typology, and syntactic description, Vol. 3 Grammatical categories and the lexicon (pp 57-149). NY: Cambridge University Press.

Tardif, T., Shatz, M., & Naigles, L. (1997). Caregiver speech and children's use of nouns versus verbs: a comparison of English, Italian, and Mandarin. *Journal of Child Language*, **24**, 535-566.

Tomasello, M**.** & Farrar, M. 1986), Joint attention and early language. *Child Development* **57**, 1454-63.

Tomasello, M. & Kruger, A.C. (1992). Joint attention on actions: Acquiring verbs in ostensive and non-ostensive contexts. *Journal of Child Language*, **19**, 1-23.

Trueswell, J.C. (1996) The role of lexical frequency in syntactic ambiguity resolution. *Journal of Memory and Language* **35**, 566-585.

Trueswell, J.C., Tannenhaus, M.K., & Garnsey, S.M. (1994). Semantic influences on parsing: use of thematic role information in syntactic ambiguity resolution. *Journal of Memory and Language*, **33**, 285-318.

Waxman, S. (1995) The development of an appreciation of specific linkages between linguistic and conceptual organization. In L. Gleitman and B. Landau (eds.), pp 229-258, *The acquisition of the lexicon,* Cambridge, MA: MIT Press.

1. Notationally, we use double quotes for the utterance of a word, italics to mention it, and single

quotes for its meaning.

2. Some question has been raised whether this generalization holds across all languages and under all child-rearing practices. For instance, Choi and Gopnik (1995) assert that because Korean mothers use twice as many verbs as nouns in speech to their children, the noun bias in early Korean production is nowhere as strong as it is in English. But as several commentators (see Gentner and Boroditsky, in press) have pointed out, these authors' reported findings in fact strengthen rather than weaken the conclusion that nouns have a privileged role in early learning, even in Korean: Despite the two-to-one input advantage of verbs over nouns in input speech, still there is a slightpredominance of nouns over verbs in the early production data (Choi and Gopnik, 1995; p. 510; Table 5a). Moreover, other word-count studies report a 4 to 1 advantage for nouns in the early vocabulary of Korean children (Au, Dapretto, and Song, 1994). Even more revealing than the spontaneous speech measures are several studies showing that children extend new nouns to further exemplars much earlier in life than they extend new verbs to new exemplars (e.g., Golinkoff, Jacquet, Hirsh-Pasek, and Nandakumar, 1996). Choi and Gopnik are not the only investigators to report variable results in cross-linguistic word counts. Tardif, Shatz, and Naigles (1997) obtained equivocal results (for English, Italian, and Mandarin) using novel procedures for categorizing first words (for example, they didn't count proper nouns as nouns). Their conclusion was that "input matters," such that linguistic and cultural factors will influence these comparative counts. This is not a controversial idea. Indeed in Gentner's classic study (1982) comparing acquisition patterns in English, Mandarin, Turkish, Kaluli, Japanese and German, she found that the *degree* of the noun advantage varies with several measureable input factors. But what stands out in this study and all others we know of is that nouns dominate the early vocabulary despite the influence of these cross-cutting factors (for a definitive review and theoretical discussion, Gentner and Boroditsky, in press).

3. We accept, and presuppose in the current work, that several interacting schemata based on prosodic and phonetic pattern-matching are at work in the infant to accomplish utterance and word segmentation (Brent and Cartwright, 1996; Fisher and Tokura, 1996; Morgan, Meier, and Newport, 1989; Saffran, Newport, and Aslin, 1996; Seidenberg and McClelland, 1989) and contribute as well to certain higher-order classifications (Brent 1997;; Gleitman, Gleitman, Landau, and Wanner, 1988; Mintz, Newport, and Bever, 1995).

4. We acknowledge that the phrase "simple words" hedges the territory to an unknown degree, depending on one's position as to the internal structure (if any) of words in the everyday vocabulary, especially that segment of it found in the earliest vocabularies. If one holds, that is, that nouns are unstructured but verbs are as a class componential, then the noun-verb disparity in learning could be approached by asserting that learning nouns was a one-step process and learning verbs a two-step process (e.g., Gentner, 1982; Landau and Gleitman, 1985, ch. 7). Experiment 2 will return to such issues but will suggest a partly different division of the lexicon.

5. Specifically, correctness of referent (as opposed to correctness of meaning)raises the percent correct of nouns from 45.03% to 52.5%, and of verbs from 15.33% to 25.9%. A repeated measures

ANOVA with the two scoring criteria as dependent measures and Lexical Class as a factor demonstrates a significant effect of lexical class between items($F = 12.29$, df=1,46 $p < .001$). In addition, there is a significant effect of the scoring method within items ($F = 18.06$, df=1, 46, $p < .001$) but no Scoring x Lexical Class interaction ($F = .62$, df = 1, 46, p ns). In short, this massively liberalized scoring, while it improves the numbers somewhat, has no qualitative effect on how one ought to think about success rate within or across lexical class for these materials.

6. Subjects were pretty well able to follow the instructions to choose within the correct class (noun or verb, depending on the target item), though there were a few exceptions (inside and there were two Finals for nouns and where and attention were two Finals for verbs). The finding that subjects chose among a smaller pool of verbs (66) than nouns (116) is not a reflection of general distinctions of frequency distribution: In the highest frequency ranges, from which just about every noun and verb selection was made by the subjects, verbs and nouns are represented about equally; it is only at lower frequency ranges that nouns start to outnumber verbs significantly (Gentner, 1981).

7. We will discuss this list effect further as the discussion progresses. As we will show, there are significant item-specific effects within as well as across lexical class on subjects' ability to glean the word meaning from context. While verb performance does not differ across the three randomly generated lists, one of them happened to have more than its share of difficult noun instances. For example, one item that achieved only one correct identification is Mommy: Subjects evidently did not see proper names as suitable responses. Another offending item is hat, whose occurrence in the taped sessions happens to be quite idiosyncratic: Four of six instances for this word showed a child using a visor (a nonstandard hat) as a carrier for a basketball, drastically lowering the correctness score for hat (see Table 1). Such effects reduced the scores for nouns on one of the lists so that this list didn't show the significant noun-verb disparity. To repeat, this is not because the verbs scores for this list were any better, rather because the scores for nouns were worse.

8. Because in this analysis we arbitrarily chose the one general verb and the one general noun that show this effect maximally, the outcome is that it seems no stronger for one class than the other . (A repeated measure ANOVA shows that a between items effect of lexical class is not significant, $F = 1.21$, df 1, 44, $p < .28$). This may not be the valid way to look at these data, however. As stated in the text, two verbs (look and play) show this sharp rise over trials, and thus account for over 27% of all false guesses, but only one noun at 11%, toy, shows either the high proportion of false alarms or the rapid rise across trials. Thus the tendency to opt for a default generality seems stronger for verbs than for nouns.

9. The problem of increasing generality as a function of increasing exposure is not inherent in all instance-based models of learning. The massed trials in our experimental situation may have artificially inflated this problem, just because they facilitated memory for the contexts that generate successive conjectures. A machinery that stores only its conjecture (with some confidence rating) on each exposure to a new word but forgets the scenario that generated this conjecture might avoid

the temptation of general solutions (for learning models with these and related properties, see Elman, 1993; Goldowsky & Newport, 1996; Newport, 1990). Because in the real-life case children will be exposed to new words sporadically rather than in uninterupted succession, they likely will be spared the curse of perfect memory in this regard.

10. Interestingly enough, toy is the default guess among the nouns and comprises 11% of the final noun guesses (Figure 4). So it is not as though this notion is difficult or insalient to the subjects, rather it is hard to identify its occurrence from contextual observation. It is selected often, but not in response to the correct input.

11. We hasten to point out that this apparent tautology is challenged vigorously by many commentators who suggest that the use of quite abstract words can be gleaned from observation of the local drift of conversational interaction even by very young children, owing to the sophisticated pragmatic skills they bring into the word-learning situation (for discussion taking various perspectives on this issue, see Pinker, 1994; Baldwin, 1991; Huttenlocher, Smiley, and Charney, 1983; Landau and Gleitman, 1985).

12. For example pronouns (e.g.,you, someone), proper names (Mommy), superordinates (thing, food), and abstractions (adoration, difference) occurred as noun conjectures, and proverbs (do), mental (try, want) and perceptual (look, listen) terms, and abstractions (coordinate, investigate) occurred as verb conjectures. Abstractions and superordinates also occurred among the mothers' most frequent word uses in Experiment 1 (e.g., think, music).

13. Probably, some subjects thought of hammer in its nominal use despite being instructed not to. A misfortune for us was that this word appeared among the 24 most frequent items for both nouns and verbs, and thus showed up twice in the merged list for the imageability test. Owing to the large number of items and the robustness of the effects, this glitch is statistically invisible.

14. The present experiments (here labelled as "Experiment 3") actually were conducted earlier, and were the impetus for the noun-verb comparison study that in the current paper is labelled Experiment 1; the order of presentation reverses this temporal order for ease of explication. The Experiment 3 sequence was designed and analyzed by three of the present authors (Lederer, Gleitman, and Gleitman).

15. It was the discovery in the present experiment (actually conducted before Experiment 1) that the differing frequencies of these verbs in maternal speech had no effect on their identifiability scores which led to our using just a standard 6 stimuli in Experiment 1.

16. The idea of increasing the video segment length, as we did in Experiment 1, was based on a suggestion from S. Pinker who offered that the lack of enough evidence for the conversatinal gist was lowering success rate in the present experiment. As we will see, he was probably right because the success rate in Experiment 1 (15%) was higher than we will report here (7%) for the shorter videos. Another factor may be the grainier black-and-white videos used here, compared to the

visually better video of Experiment 1.

17. . For these analyses, the percent correct data were submitted to a log transformation to correct for scaling problems in the data. Duncan's Multiple Range Test was also calculated to correct for multiple comparisons.

18. See Resnik, 1995, for an elegant computational model of selectional constraints that considers its relevance to lexical acquisition.

19. Dowty (1991) contains a major linguistic statement of such relations between conceptual prominence and subjecthood; see also Talmy, 1978 for a related picture. For on-line effects of these and related factors in sentence comprehension, see Trueswell, Tanenhaus and Garnsey, 1994. A number of studies show that even where the predicate itself is symmetrical and where (therefore) the ordering of its nominals "should not matter," there are powerful effects of prominence on subjecthood (Talmy, 1978; Gati and Tversky, 1984; Gentner and Rattermann, 1991; Gleitman, Gleitman, Miller, and Ostrin, 1994). For discussion and evidence from learning, see Fisher et al., 1994 and Miller, 1998. All the same, the errorfulness of a procedure that puts too much faith in this kind of information is evident in the *call* examples (Figure 7) in which such animates as *Daddy, Markie,* and *Grandma* show up in complement position.

20. Despite this absolute improvement in success rate when the videotaped scenes were added to the information base of Condition 5, the level of improvement did not reach statistical significance (p = .10). Likely this is a ceiling effect: The distributional evidence, syntactic and selectional, is evidently so determinative of verb semantics that there is not so much left for observation to do. Putting this another way, the selctional restrictions (the nouns in their thematic role positions) are informative in much the same way as are the visible entities -- Markie, the phone, etc. -- interacting in the extralinguistic world. Recall also that success rate for the video-only Experiment 1 was 15%.for verbs using the better video and longer segments of Experiment 1, and only 7% in Condition 1 of the present experiment, so the 90% success rate in the present condition, high as it is, probably is an underestimate of how easy verb identification is, given multiple sources of evidence.

Table 1
Correct Identification of Target
Final Conjecture in Experiment 1
(targets listed from most to least frequent)

Noun Targets	Final Conjecture	Verb Targets	Final Conjecture
piggy	89.3	go	3.6
ball	78.6	do	3.6
mommy	3.6	put	35.7
hat	28.6	come	75.0
elephant	89.3	want	3.6
plane	100	see	10.7
bag	85.7	look	42.9
kiss	7.1	get	7.1
toy	25.0	turn	3.6
drum	89.3	play	21.4
people	39.3	hammer	14.3
nose	67.9	have	0
hole	57.1	push	42.9
daddy	3.6	say	0
music	39.3	throw	85.7
hand	14.3	pop	0
tail	25.0	like	0
hammer	71.4	stand	3.6
thing	3.6	think	0
camera	46.4	know	0
peg	10.7	make	0
pilot	3.6	wait	3.6
shoes	3.6	fell	10.7
swing	96.4	love	0
Mean	**44.9**	**Mean**	**15.3**

Table 1

Table 2
Most Frequent Response
on Trial 1 and Final Conjecture
(Targets listed from most to least frequent)
(Correct responses indicated in boldface)

Noun Targets	Trial 1 Response	Final Response	Verb Targets	Trial 1 Response	Final Response
piggy	**piggy**	**piggy**	go	put	hit
ball	**ball**	**ball**	do	clap	look
mommy	drum	toy	put	come	**put**
hat	head	**hat**	come	**come**	**come**
elephant	**elephant**	**elephant**	want	turn	play
plane	**plane**	**plane**	see	get	look
bag	**bag**	**bag**	look	**look**	**look**
kiss	food	mouth	get	catch	hold
toy	drum	**toy**	turn	look	play
drum	**drum**	**drum**	play	**play**	**play**
people	plane	**people**	hammer	hit	put
nose	elephant	**nose**	have	play	play
hole	**hole**	**hole**	push	**push**	**push**
daddy	person	phone	say	swing	push
music	drum	drum	throw	catch	**throw**
hand	elephant	toy	pop	hit	get
tail	bag	**tail**	like	drive	look
hammer	knee	**hammer**	stand	stop	stop
thing	blanket	toy	think	tickle	look
camera	dog	**camera**	know	play	look
peg	piece	nail	make	hit	put
pilot	person	people	wait	stop	open
shoes	**shoes**	feet	fell	look/hit	look
swing	**swing**	**swing**	love	hit	play
# most frequent response correct	9	15	# most frequent response correct	4	6

Table 2

Table 3
Imageability Rating Scores[*]
Experiment 2
(targets listed from most to least frequent)

Noun Targets	Noun Rating	N & V Rating	Verb Targets	Verb Rating	N & V Rating
piggy	5.0	6.26	go	4.2	3.37
ball	6.8	6.52	do	3.7	2.07
mommy	6.4	6.44	put	4.2	3.04
hat	6.6	6.33	come	4.5	3.26
elephant	6.9	6.41	want	2.3	2.93
plane	6.5	6.63	see	4.6	4.07
bag	6.6	6.41	look	4.8	3.59
kiss	4.9	5.82	get	3.8	2.59
toy	4.8	6.04	turn	5.8	4.37
drum	6.5	6.26	play	5.2	4.37
people	4.8	6.26	hammer	6.7	5.56
nose	6.1	6.90	have	2.7	2.93
hole	5.2	5.82	push	5.5	4.44
daddy	6.4	6.41	say	4.0	4.22
music	4.4	4.93	throw	6.0	4.63
hand	6.8	6.07	pop	4.6	4.56
tail	5.3	6.07	like	3.3	2.82
hammer	6.6	6.37	stand	5.8	4.52
thing	2.9	2.52	think	4.2	3.07
camera	6.7	6.52	know	2.7	2.22
peg	4.4	5.67	make	3.8	2.96
pilot	5.2	6.07	wait	3.1	3.41
shoes	6.5	6.41	fell	6.2	3.33
swing	5.7	6.26	love	4.6	3.78
Mean	**5.75**	**6.08**	**Mean**	**4.43**	**3.59**

[*]1=least imageable, 7= most imageable; Noun Rating--Ss rated only nouns; Verb Rating--Ss rated only verbs; N & V Rating--Ss rated nouns and verbs together.

Table 3

Table 4
Design for Experiment 3

Condition	Source of Information
1 Cross-situational Observation	Videoclips
2: Noun Co-occurrence	Alphabetical lists of nouns for each sentence
3: Observation and Noun Context	Videoclips & list of nouns (Cond 1 + 2)
4. Syntactic Frame	Sentences with nonsense nouns and verbs
5 Syntax and Selection	Sentences with only verb unidentified
6 Full Information	Sentences & videoclips (Cond 1 + 5)

Table 4

Table 5
Correct Identification (in %) of Target Verb in Experiment 3
(Targets listed most to least frequent)

TARGET	Cross-Situational Observation Cond 1	Noun Co-Occurrence Cond 2	Observation and Noun Context Cond 3	Syntactic Frames Cond 4	Syntax + Selection Cond 5	Full Info Cond 6
go	0	20	15	60	65	100
see	0	25	5	90	95	100
come	0	0	15	40	60	100
say	0	35	20	85	90	100
do	0	5	0	15	35	100
put	5	15	50	75	95	100
get	10	0	30	40	65	100
look	25	45	30	95	100	100
want	0	30	55	90	100	100
have	0	0	30	30	45	100
know	0	0	0	90	100	100
like	0	0	0	80	90	100
think	0	0	0	90	90	100
take	15	35	20	60	85	100
find	10	5	35	45	65	80
play	5	45	75	50	85	100
push	50	30	70	15	90	90
show	0	10	50	90	100	100
sit	5	5	25	15	55	100
catch	35	5	15	10	45	40
call	5	30	50	20	65	60
make	0	35	50	25	70	80
eat	15	0	5	5	35	30
pull	5	20	50	25	85	90
Mean	**7.7**	**16.5**	**29.0**	**51.7**	**75.4**	**90.4**

Table 5

Table 6
Incremental Pairwise Comparisons**
of Sources of Learning Information in Experiment 3

Condition	Mean % Correct	Pairwise Comparisons T	p value
1: Cross-situational Observation	7.7		
		2.16	.10
2: Noun Co-occurrence	16.5		
		2.13	.05
3: Observation & Noun Context	29.0		
		2.77	.01
4. Syntactic Frame	51.7		
		3.15	.01
5: Syntax and Selection	75.4		
		2.36	.10
6: Full Information	90.4		

**Duncan Multiple Range Test performed to correct for multiple comparisons (df=23).

Figure Legends

FIGURE 1: Vocabulary composition as a function of vocabulary size on the MacArthur CDI Toddler Scale, reprinted from Bates et al. (1995). The dotted lines represent the proportion of items from each lexical class sampled in the full Toddler Scale.

FIGURE 2: Mean identification score of target words by trial in Experiment 1.

FIGURE 3: The rise of *look* and *toy* across trials.

FIGURE 4. These cartoons approximate four of the seven (approximately 20 seconds long) video segments for the verb *call*. In Condition 1, subjects saw this video alone, in Condition 3 subjects saw the video accompanied by the matching noun list of Figure 5, and in Condition 6 subjects saw the video accompanied by the matching sentences as shown in Figure 7.

FIGURE 5: Noun-within-sentence presentations for Conditions 2 and 3, for the verb *call*. In Condition 2, subjects saw only these lists. In Condition 3, each was accompanied by its matching video (Figure 4).

FIGURE 6: Sentence frames with the verb represented by a (capitalized) nonsense item and the nouns representated by lower case nonse, for the verb *call*, as presented in Condition 4.

FIGURE 7: The maternal sentences with a nonsense word (capitalized) substituted for the verb *call*. In Condition 5, subjects saw these lists only. In Condition 6, they were accompanied by the matching video (Figure 4).

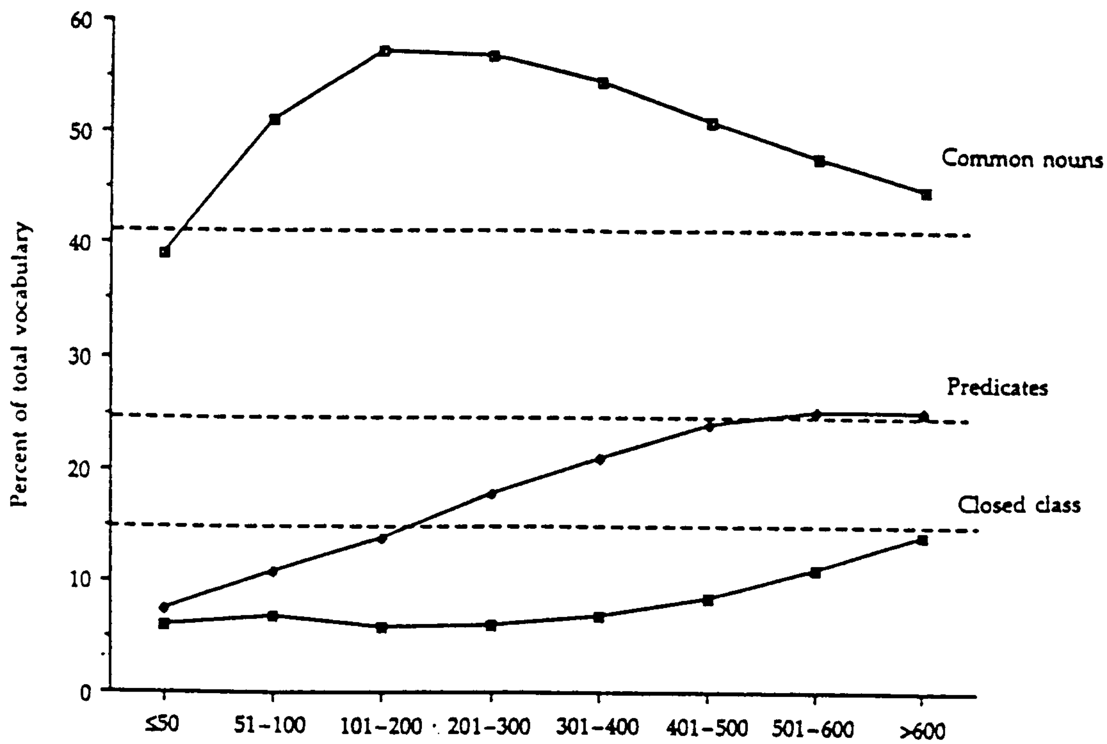
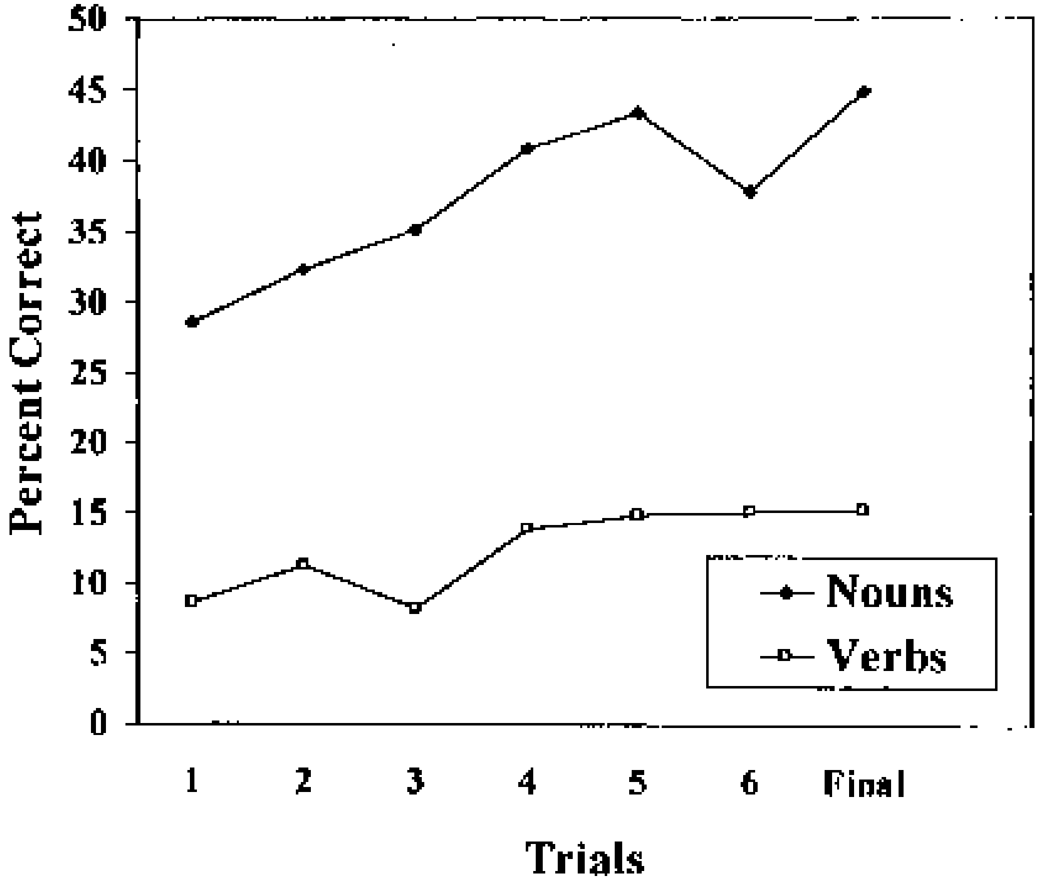
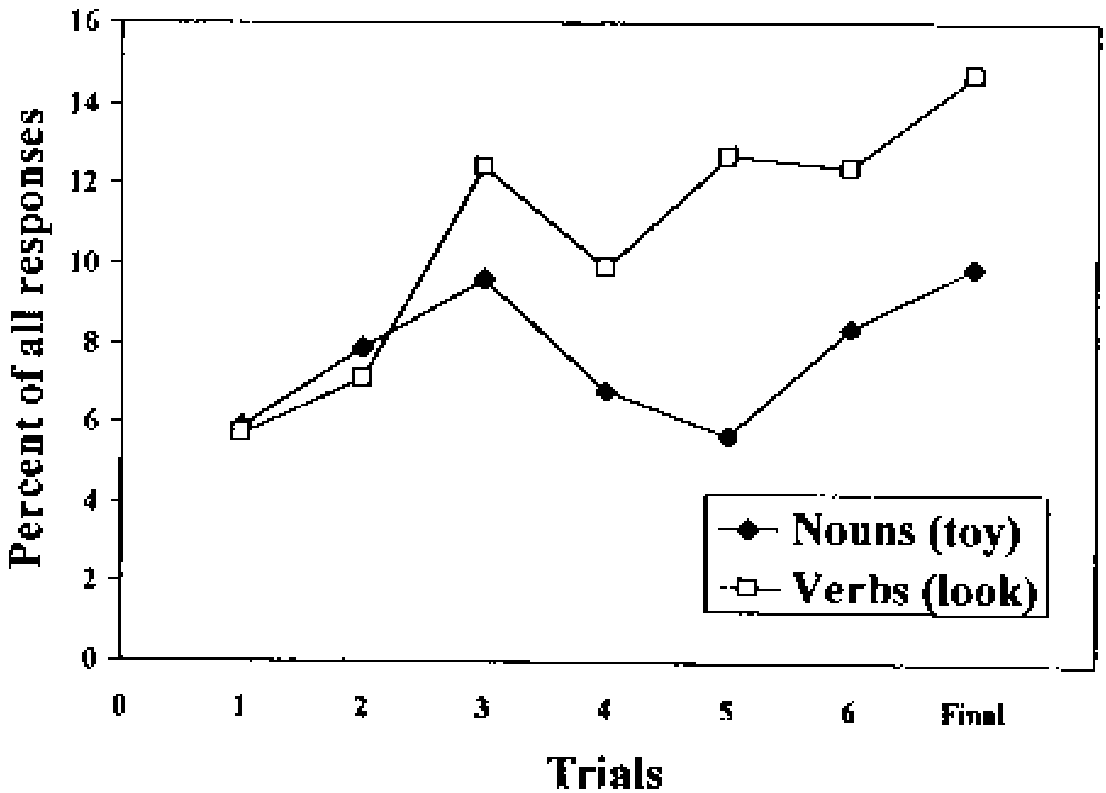

60
50
40
30
20
10
0
Percent of total vocabulary
Common nouns
Predicates
Closed class
≤50
51-100
101-200
201-300
301-400
401-500
501-600
>600

Figure 1



50
45
40
35
30
25
20
15
10
5
0
Percent Correct
Nouns
Verbs
1
2
3
4
5
6
Final
Trials

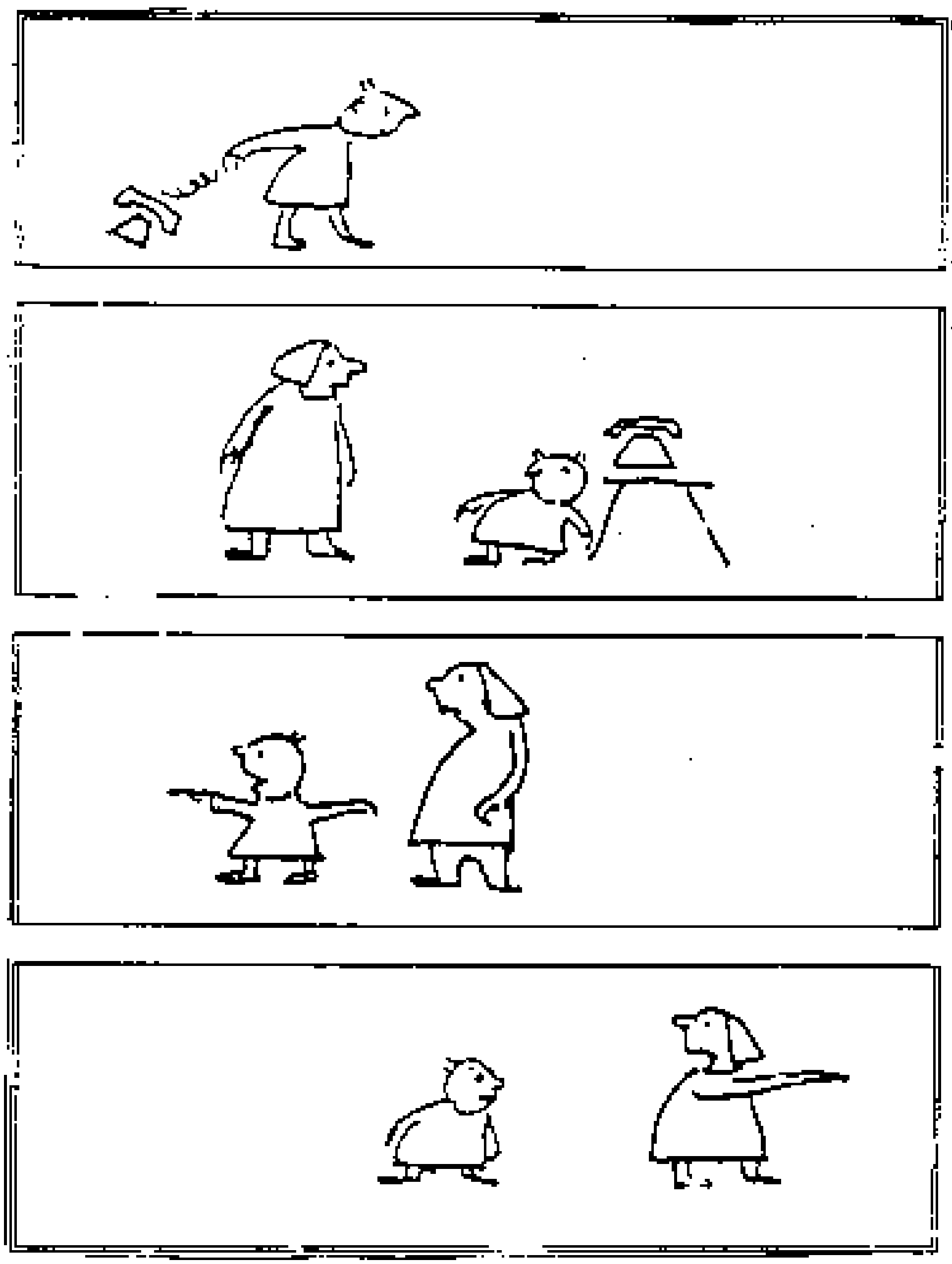
Figure 2



16
14
12
10
8
6
4
2
0
Percent of all responses
0
1
2
3
4
5
6
Final
Trials
Nouns (toy)
Verbs (look)

Figure 3

"What is *Gorp*?"



etc.

Figure 4

Mystery verb:	Co-occurring nouns in alphabetical order:
GORP means.....?	gramma, you Daddy, Daddy Daddy, you I, Markie Markie, phone, you Mark Mark.

Figure 5

Mystery Verb:	Syntactic Frames:
GORP means.....?	Why don't ver GORP telfa? GORP wastorn, GORP wastorn. Ver gonna GORP wastorn? Mek gonna GORP litch. Can ver GORP litch on the fulgar? GORP litch. GORP litch.

Figure 6

Fig 6

Mystery Verb:	Syntax and Selectional Information:
GORP means.....?	Why don't you GORP gramma? GORP Daddy, GORP Daddy. You gonna GORP Daddy? I'm gonna GORP Markie. Can you GORP Markie on the phone? GORP Mark. GORP Mark.

Figure 7