DOCUMENT RESUME

ED 335 884 FL 019 337

AUTHOR Gleitman, Lila R.
TITLE The Structural Sources of Verb Meaning.
PUB DATE Apr 89
NOTE 50p.; In: Papers and Reports on Child Language Development, Number 28, p1-48, Aug 1989. For the proceedings, see FL 019 336.
PUB TYPE Speeches/Conference Papers (150)

EDRS PRICE MF01/PC02 Plus Postage.
DESCRIPTORS *Child Language; Cognitive Mapping; *English; *Language Acquisition; Language Research; Linguistic Theory; *Semantics; Structural Analysis (Linguistics); *Verbs; *Vocabulary Development

ABSTRACT

A discussion of English native-language vocabulary acquisition in children takes a closer look at the assumption that vocabulary is learned by common association of word with event, focusing on the acquisition of verb meanings. The intuitive power of the view that words are learned by noticing real-world contingencies for their use is acknowledged, but it is pointed out that such mapping, unaided, is in principle insufficiently constrained to explain how the child maps verbs (as phonological objects) with their meanings. The solution offered is that semantically relevant information in the syntactic structures can rescue observational learning from experiential pitfalls. Evidence is offered that children deduce meanings from their knowledge of structural-semantic relations. Limitations in data, need for further information about cross-linguistic correspondences, and problems occurring in the analysis are briefly addressed. A 59-item bibliography is included. (MSE)

PRCLD 28 (1994)
Keynote Address

The structural sources of verb meaning*

Lila Gleitman

> If we will observe how children learn languages, we will find that, to make them understand what the names of simple ideas or substances stand for, people ordinarily show them the thing whereof 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'. (Locke, 1690, Book 3.IX.9)

Is vocabulary acquisition as straightforward as Locke supposes? Three hundred years after the publication of the Essay on Human Understanding, Locke's is still the dominant position on this topic for the very good reason that common sense insists that he was right: Word meanings are learned by noticing the real-world contingencies for their use. For instance, it seems obvious to the point of banality that the verb pronounced /run/ is selected as the item that means 'run' because this is the verb that occurs most reliably in the presence of running-events.

Or is it? Who has ever looked to see? One trouble with questions whose answers are self-evident is that investigators rarely collect the evidence to see if they pan out in practice.

Since this occasion of a keynote address is a serious one, I certainly am not going to try to defeat the obviously correct idea that a crucial source of evidence for learning word meanings is observation of the environmental conditions for their use. I believe, however, that what is correct about such a position is by no means obvious, and therefore deserves serious study rather than acceptance as a background fact in our field.

I'll limit the discussion to the topic of acquiring verb meanings, because this is where I and my colleagues have some experimental evidence to offer in support of the position I want to adopt. Even within this subtopic, to begin at all I will have to make critical assumptions about some heady issues which deserve study in their own right. Particularly, I will not ask where the concepts that verbs encode come from in the first place, for example, how the child comes to conceive of such notions as 'run' (or 'think' or 'chase'). I want to look at the learner at a stage when he or she can entertain such ideas, however this stage was arrived at. Second, I reserve for later discussion the question of how the child determines which word in the heard sentence is the verb -- that it is the phonological object /run/, not /horse/ or /marathoner/, that is to be mapped

onto the action concept.

The topic that remains seems a very small one: How does the learner decide which particular phonological object corresponds to which particular verb concept, just Locke's topic. But I'll try to convince you that this question is harder than it looks. For one thing, matching the meanings to their sounds is the one part of acquisition that can't have any very direct "innate" support; this is because the concept 'run' isn't paired with the sound /run/ in Greek or Urdu, so the relation must be learned by raw exposure to a specific language. For another thing, and as I'll try to convince you today, it's not clear at all that the required pairings are available to learners from their ambiant experience of words and the world.

In the first half of this talk, I'll try to set out some of the factors that pose challenges to the idea that children can induce the word meanings from their contexts in the sense Locke and his descendents in developmental psycholinguistics seem to have in mind. In this discussion, I will allude repeatedly to the work and theorizing of Steve Pinker, because he seems to me to be the most serious and acute modern interpreter of ideas akin to Locke's in relevant regards. Then, in response to these challenges to the theory of learning by observation, I will sketch a revised position laid out by Landau and Gleitman (1985), illustrating it with some recent experimental evidence from our laboratory. The idea here is that, to a very considerable extent, children deduce the verb meanings by considering their syntactic privileges of occurrence. They must do so, because there is not enough information in the whole world to learn the meaning of even simple verbs.

Part I: Some difficulties of learning by observation

Locke's idea: Differences in experience should yield differences in meanings

At peril of carricaturing Locke -- but who doesn't? -- I select him as one who argued for a rather direct relation between knowledge and the experience of the senses. He frequently used the example of individuals born without sight as a testing ground for such a position. According to Locke, sighted and blind people ought both to be able to learn the meanings of such words as statue and feel and sweet, but the blind ought to be unable to acquire picture and see and red, for these concepts are primitive (i.e., not derivable from other concepts) or derivable from primitives that are available only to the eye.

Barbara Landau and I were directly inspired by Locke to study the acquisition of these vision-related terms by blind babies (Landau and Gleitman, 1985). As our studies evolved, we

realized that exactly the same conceptual issues about learning arise for sighted vocabulary learners as for blind ones, so I will move on to discussion of such normally endowed children. The blind population, which I discuss first, is perhaps special only as the biographical point of origin of our own thinking but I suspect that, for you listeners too, it will serve to dramatize some issues which seem less startling in the ordinary case. These have to do with how resistant the word-learning function is to the evidence of the senses.

Landau and I were astonished to discover how much alike were the representations of vision-related terms by blind and sighted children at age 2 1/2 or so, despite what would appear to be radical differences in their observational opportunities. For instance, all these babies showed by their comprehension performances that they took <u>look</u> and <u>see</u> as terms of perception, distinct from such contact terms as <u>touch</u>. As an example of this, a blind child told to "Touch but don't look at..." a table would merely bang or tap it. Whereas if told "Now you can look at it" she explored all its surfaces systematically with her hands. Moreover, she understood <u>look</u> to be the active (or exploratory) and <u>see</u> the stative (or achievement) term in this pair. Just as surprising, blind children as well as sighted children understood that <u>green</u> was an attribute predicable only of physical objects (they asserted that ideas could not in principle be green while cows might be, for all they knew). Thus the first principle that a theory of observational learning must be subtle enough to capture is that

(i) The same semantic generalizations can be acquired in relative indifference to differing environmental experience, if the notion "experience" is cast in sensory-perceptual terms.

<u>Can word-to-world pairings in the input account for the child's semantic conjectures?</u>

While we found the surprising result that blind children shared much knowledge about vision-related terms with their sighted peers, we also achieved the unsurprising result that there were some differences in how these two populations understood these terms to refer to their own perceptions: Blind children think that <u>look</u> and <u>see</u> describe their own <u>haptic</u> perceptions while sighted children think these same words describe their own <u>visual</u> perceptions. Thus blindfolded sighted children of 3 years look skyward if told to "Look up!" but a blind child of the same age holds its head immobile and searches the space

above in response to the same command (see Figures 1 and 2).[1]

This outcome is of just the sort that is subject to "obvious" explanations involving the extralinguistic contexts of use. We reasoned (as does everyone to whom one presents this set of facts): 'Obviously,' a blind child's caretaker will use the terms look and see intending the child to perceive in whatever ways her sensorium makes available. And since the blind child's way of discovering the nature of objects is by exploring them manually, the caretaker will surely use look and see to this child only when an object is near enough to explore manually. That is, the caretaker should say "Look at this boot" to her blind baby only if a boot is nearby, ready to be explored manually. The contexts of use for these words thus should include -- among many other properties -- conversationally pertinent objects that are near at hand. Had the caretaker instead rattled a boot noisily by the child's ear whenever she said "Look at this boot", the learner would have surmised that look means 'listen'.

So here we have a straightforward prediction from the environment of use to the formation of a semantic conjecture: By hypothesis, the blind learner surmises that look involves haptic exploration because it is that verb which is used most reliably in contexts in which haptic exploration is possible and pertinent to the adult/child discourse. Landau and I decided to test that prediction to see if it was as true as it was obvious.

To do so we examined videotapes of a mother and her blind child recorded in the period before the child uttered any vision-related words or indeed any verbs at all (that is to say, during the learning period for these words), coding all verb uses according to whether they occurred when an object pertinent to the conversation (a) was NEAR enough to the child for her to explore it manually, i.e., within arm's reach, (b) was FARther away than that, or (c) when there was NO such pertinent OBJECT. We hypothesized that look and see were the verbs used most reliably in the NEAR condition accounting for why the child had

---

[1] A related difference holds for the color words. Sighted children of four and five map the color words onto observed hues in the world while blind children ask for help. Perhaps they think the property is stipulative. Asked "Why are the flowers in the woods pink?" one blind child responded "Because we name them pink!" (Landau, personal communication). They know these are attributes predicable only of physical objects (they say that an idea can't be green because "it's only in your head") but they don't know what the real-world dimension may be. Interestingly, they avoid some choices that their extralinguistic experience appears to make available, e.g., that color terms refer to sizes of objects (Landau and Gleitman, ibid, ch. 8).

Figure 1: A blindfolded sighted child's response to the command "Look up!" (from Landau and Gleitman, 1985)

Figure 2: A blind child's response to the command "Look up!" (from Landau and Gleitman, 1985)

assigned them the meanings 'explore/apprehend manually' (while other verbs would be used less often in this condition, and so would not be assigned a haptic component in their meanings).

The results are shown in Table 1. They fail to account for the child's haptic interpretation of look and see. Put and give and hold are the verbs used most reliably (over 95% of the time) under the NEAR condition while look (73%) and especially see (39%) are not as reliably associated with this condition. We can conclude that

(ii) If representations of the environmental contexts are the basis for the semantic conjectures, these can't can't be just the simplest and most obvious representations of those contexts that one can think of.

It is worth pointing out before leaving this topic that the analysis of Table 1 cannot be written off as of some environmental property that is hopelessly irrelevant to the child's analysis of events (though it is doubtless too simple, a fact to which I will return directly). For as it stands, this analysis extracts and explains important distinctions among verbs of physical motion that are in other respects semantically close, such as give vs get. The child is apparently told, sensibly enough, to give what she has in hand (this verb is used in the NEAR condition 97% of the time) but to GET what she doesn't have (the relevant NEAR percentage for this verb is 45%).

Latitude of the hypothesis space

Generalization (ii) brings me closer to topics I want to concentrate on today. Notice that the conclusion drawn was very weak -- not that it wasn't the contexts that led to the learning, but rather that the idea of "real-world context," to succeed, must be a good deal more subtle than we (and others) originally supposed. That is, the response to the findings shown in Table 1 is usually, and perhaps should be:

"Oh, but the contextual analysis you imposed was so feeble. Showing that it failed is only showing the failure of Landau and Gleitman's imagination. The child surely imposes a richer analysis on the situation than that, and the only analysis relevant to the hypotheses under test is the one that the child herself imposes."

Fair enough. We limited the child to observing some perceptually obvious features of the situation, features that the infancy literature tells us are available even to babies. In other words, our aim was to see how far some small and independently documented set of observational primitives could get the learner in extracting simple meaning features for assignments to the verbs. These were that the world is

Verb	Proportion used in contexts: In hand or near	Far	No object	Total number considered[a]
Perceptual verbs				
Look	.73	.09	.18	34
See	.39	.56	.05	18
Other perceptual	.56	.44	.00	17
Nonperceptual verbs				
Come	.05	.32	.63	19
Get	.50	.25	.25	27
Give	.97	.03	.00	21
Go	.52	.24	.24	20
Have	.53	.47	.00	11
Hold	1.00	.00	.00	10
Play	.70	.00	.30	10
Put	.97	.00	.03	61
Say	.43	.07	.50	28

a. These total to N=276, the number of utterances containing the common verbs (10 or more occurrences in the maternal corpus). The remaining 389 were discarded in this and following analyses, including 183 instances of *be* and 186 instances of rare verbs (fewer than 10 occurrences).

Table 1: Situational contexts for the common verbs used by the blind child's mother during the learning period (from Landau and Gleitman, 1985)

populated with objects which endure over time (Spelke, 1982), and which move relative to each other (Lasky & Gogol, 1978) and with respect to the positions of the child's own body (Acredolo and Evans, 1980; Field, 1976). These assumptions put the child in a position to conceive of the situation as one of objects--in this case, objects whose noun names are known to the child--moving (as described by the verb) between sources and goals. For example, for give the object moves from NEAR as action begins to FAR when it ends, and in get the object goes from FAR to NEAR.[2]

It can hardly be denied, in light of the infancy evidence, that youngsters do represent situations in terms of the positions and motions of pertinent objects. What is surely false, however, is that such categories are exhaustive amongst the child's extralinguistic analyses. Infants come richly prepared with means for picking up information about what is going on in their environment -- looking, listening, feeling, tasting, and smelling; in fact these different sensory routes appear to be precoordinated for obtaining information about the world (Spelke, 1979). To take a few central examples, infants perceive the world as furnished with objects which are unitary, bounded, and persist over time and space (see Gibson and Spelke, 1983), and which cannot occupy two places at one time (Baillargeon, Spelke, and Wasserman, 1985). They distinguish among the varying properties of objects, e.g., their rigidity or elasticity (Gibson and Walker, 1984), their size (Golinkoff et al, 1984a), their colors (Bornstein, 1975), whether they are moving or stationary (Ball and Vurpillot, 1976), their positions and motions relative to the child observer (Field, 1976), their animacy (Golinkoff et al, 1984b) and even their numerosity (Starkey, Gelman, and Spelke, 1983). If you think there's something that infants can't or won't notice, look in the next issue of Developmental Psychology and you will probably discover that someone proved they can.

Now that I have acknowledged something of the richness of infant perception, why not let the learner recruit this considerable armamentarium for the sake of acquiring a verb vocabulary? That is, why not assume that the child encodes the situation not only in the restricted terms that yield Table 1, but in myriad other ways? For instance, over the discourse as a whole, probably the mother has different aims in mind when she tells the child to "look at" some object than when she tells her to

---

[2] We hasten to say such an analysis can succeed at all only if the child can determine the discourse addressee. This assumption is plausible because (1) in these transcripts, at least, the mother's speech is over 95% about the "here and now" and (2) in over 90% of instances, the addressee is the child herself.

"hold" or "give" it. The child could code the perceptual world for these perceived aims and enter these properties as aspects of the words' meanings. But also the mother may be angry or distant or lying down or eating lunch and the object in motion may be furry or alive or large or slimy or hot, and the child may code for these properties of the situation as well, entering them too as facets of the words' meanings.

The problems implicit in such an expansion of the representational vocabulary should be familiar from the literature on syntax acquisition: The trouble is that an observer who notices <u>everything</u> can learn <u>nothing</u>, for there is no end of categories known and constructable to describe a situation.[3]

Indeed, not only learnability theorists but all syntacticians in the generative tradition appeal to the desireability of "narrowing the hypothesis space" lest the child be so overwhelmed with representational options and data-manipulative capacity as to be lost in thought forever. At least, learning of syntax could not be as rapid and uniform as it appears to be, unless the child were subject to highly restrictive principles of Universal Grammar, which rein in her hypotheses. As one famous example, the learner is said to assume that all syntactic generalizations are structure-dependent rather than serial-order dependent (Chomsky, 1975; see also Crain and Fodor, in press). In fact, Universal Grammar is said to be as constrained as it is owing to the child's requirement that this be so.

I put it to you: Are these observations about the difficulties of learning when the hypothesis space is vast no less true of word learning than of syntax? In the domain of vocabulary acquisition as much as that of syntax acquisition, there is remarkable efficiency and systematicity of learning across individuals (and, as the blind children show, across learning environments): The rapidity and accuracy of vocabulary acquisition are jewels in the crown of rationalistically ori-

---

[3] As so often, Chomsky (1982) sets the problem with great clarity: "...The claim we're making about primitive notions is that if data were presented in such a way that these primitives couldn't be applied to it directly, prelinguistically before you have a grammar, then language couldn't be learnt...And the more unrealistic it is to think of concepts as having those properties, the more unrealistic it is to regard them as primitives ...We have to assume that there are some prelinguistic notions which can pick out pieces of the world, say elements of this meaning and of this sound." The analysis of Table 1 is an attempt to see how far some small and independently documented set of observational primitives could get the learner in extracting a simple meaning feature (<u>'haptic'</u>) for assignment to certain verbs.

ented developmental psycholinguistics (see particularly Carey, 1982). So just as in the case of syntax, we have initial grounds for claiming that a limit on the hypothesis space must be a critical source of sameness in the learning function. Bolstering the same view, languages seem to be as alike in their elementary vocabularies as they are in their syntactic devices (see for example Talmy, 1975; 1985). But surprisingly enough, all the telling arguments, invoked for syntax, to restrict the interpretation of the input -- that is, constraints on representations -- that are to explain these samenesses in form, content, and learning functions, are thrown out the window in most theorizing about the lexicon. There it is usually maintained that the child considers many complex, varying, cross-cutting, subtle conjectures about the scenes and events in view so as to arrive at the right answers, comparing and contrasting possibilities across many events, properties, discourse settings, and so forth. In other words, testing and manipulating an exceedingly broad and free-ranging hypothesis space.

A very few investigators have been responsive to the issues here. Pinker (1987), in a direct and useful discussion of the requirement to limit the space of observables that a learner will consider in matching the event to the unknown verb, writes as follows:

> Verbs' definitions are organized around a surprisingly small number of elements: "The Main Event", that is, a state or motion; the path, direction, or location of an object, either literal spatial location or some analogue of it in a nonspatial semantic field; causation; manner; a restricted set of the properties of a theme or actor; temporal distribution (aspect and phase); purpose; coreferentiality of participants in an event; truth value (polarity and factivity); and a handful of others.
>
> (1987, p. 54)

It is an open question whether Pinker's proposed list is narrow enough to meet the requirement for a realistic set of primitives upon which a verb-learning procedure can operate. Are purposes, truth values, causes, not to speak of "analogues of spatial location in nonspatial semantic fields" really primitives that inhere in the observations themselves? It seems to me highly unlikely that any choice of perceptual constraints will be restrictive enough to delimit the analyses a child performs in reaction to each event/ verb pair. Of course I'm not suggesting that there aren't principles of perception that are restrictive and highly structured (God forbid!). But they are likely not restrictive enough to account for vocabulary acquisition. How could they be? Perception has to be rich enough to keep the babies from falling off cliffs and mistaking distant tigers for nearby pussycats lest they all disappear from

the face of the earth before learning the verb meanings.

However, the richness of perception is not the only, or even the major, problem faced by a hypothetical learner who tries to acquire verb-meanings from observation. The more difficult problem is that even the homeliest and simplest verbs, though they refer to events perceivable, encode also the unobservable present interests, purposes, beliefs, and perspectives of the speaker. I turn now to this class of problems.

Perspectives on events

Consider the learning of simple motion verbs, such as push or move. In a satisfying proportion of the times that care-caretakers say something like "George pushes the truck," George can be observed to be pushing the truck. But unless George is a hopeless incompetent, every time he pushes the truck, the truck will move. So a verb used by the caretaker to describe this event may represent one of these ideas ('push') or the other ('move').

Moreover, every real event of the pushy sort necessarily includes, in addition to the thrust and goal, various values of trajectory, rate, and so forth, so that such ideas as 'slide,' 'rumble,' 'roll,' 'crawl,' and so on, are also relevant interpretations of a new verb then uttered. What is left open by the observation is whether that verb represents any or all of these manner differences: no, in the case of push, but yes in the case of roll or rumble.

Note that the manner elements just mentioned do fall within the range encoded by verbs in many languages (Talmy, 1985) and are on the narrowed list of perceptual properties suggested by Pinker (1987). I leave aside various other interpretations often called "less salient" (i.e., I ignore more general consideration of the "stimulus-free" character of language use; see Chomsky, 1959), especially the countless zany interpretations of this event that could be drawn by worried philosophers.[4]

---

[4] Jerry Fodor has suggested to me, maybe seriously, that these problems go away because the caretaker and child are in cahoots, and they are mind-readers. They are so attuned in discourse, being creatures of exactly the same sort, that the child zaps onto exactly the characteristics of the situation that the mother, just then, has in mind to express (see Bruner, 1974/5 for a story about how the attentional conspiracy is set up by mother and child, and Slobin, 1977, for a related account of the conversational environment). A related position is maintained by Pinker (citing Keenan) about situations the learner might select as learning opportunities; in the case

It is plausible that these ambiguities are eliminated by looking at a verb's uses <u>across</u> situations (see again Pinker, 1987). There will eventually be some instance of moving called /<u>push</u>/ in which the truck is moving rapidly, eliminating 'crawl' as a conjecture about the meaning of this item, etc. By a process of cross comparison and elimination, each verb may eventually be distinguishable. The worry is only that the burden of hypothesis testing is becoming ominous as the comparison set (of verbs, properties, and scenes) enlarges.[5]

Difficult problems can be solved. But impossible ones are harder. Consider such verbs as <u>flee</u> and <u>chase</u>, <u>buy</u> and <u>sell</u>, <u>win</u> and <u>beat</u>, <u>give</u> and <u>get</u>, and so on. Such pairs are common in the design of verb lexicons. Each pair alludes to a single kind of event: Whenever the hounds are chasing the fox, the fox is fleeing from the hounds. If some hounds are racing, even with evil intentions, toward a brave fox who holds its ground, they can't be said to be chasing him. Chasing implies fleeing, necessarily. If the child selects a verb from the stream of speech accompanying such a scene, how then is she to decide whether it means 'chase' or 'flee'?

Pinker is discussing, the child is to discover the property <u>subject</u> from its semantic/pragmatic environmental correlates:

> The semantic properties of subject hold only in basic sentences: roughly, those that are simple, active, affirmative, declarative, pragmatically neutral and minimally presuppositional...The parents...or the child might filter out nonbasic sentences from the input using various contextual or phonological diagnostics of nonbasicness such as special intonation, extra marking on the verb, presuppositions set up by the preceding discourse or the context, nonlinguistic signals of the interrogative or negative illocutionary force of an utterance, and so on. (Pinker, 1984, pp 46/7).

Note again the number and nontransparency of the experiential analyses necessary within this perspective.

[5] I may well be granting too much here. After all, touching, and even breathing and existing are going on in the presence of all moving and pushing events. So it's probably not true that a unique interpretation of verbs from scenes can ever be extracted, whatever the ornateness of the scene-storage and manipulation procedures may be. Not at least without invoking notions of "salience" which is likely just substitution of unknowns for unknowns.

Such examples are thrusts to the heart of the observational learning, hypothesis. As Pinker (1987, p. 54) acknowledges, "Basically, we need to show that the child is capable of entertaining as a hypothesis any possible verb meaning, and that he or she is capable to eliminating any incorrect hypothesis as a result of observing how the verb is used across situations." But chase and flee (and a host of similar pairs) are relevant used in all and only the same situations. It follows that cannot be shown that the child is capable of eliminating the incorrect hypotheses by cross-situational observation.

I think the problem is that words don't describe events simpliciter. If that's all words did, we wouldn't have to talk. We could just point to what's happening, grunting all the while. But instead, or in addition, the verbs seem to describe specific perspectives taken on those events by the speaker, perspectives which are not "in the events" in any direct way. How far are we to give the learner leave to divine the intents of his elders as to these perspectives? Are they talking of hounds acting with respect to foxes, or of foxes with respect to hounds?

Speaking more generally, since verbs represent not only events but the intents, beliefs, and perspectives of the speakers on those events, the meanings of the verbs can't be extracted solely by observing the events.

The subset problem

A related problem has to do with the level of specificity at which the speaker, by the words he chooses, refers to the world. Consider the homely little objects in the world, the pencils, the ducks, the spoons. All these objects are supplied with more than one name in a language, e.g., animal, duck, Donald Duck. I expect that the adult speaker has little difficulty in selecting the level of specificity he or she wants to convey and so can choose the correct lexical item to utter in each case. And indeed, the learner may be richly pre-equipped perceptually and conceptually so as to be able to interpret scenes at these various levels of abstraction, and to construct conceptual taxonomies (Keil, 1979). But as usual this very latitude adds to the mystery of vocabulary acquisition, for how is the child to know the level encoded by the as yet unknown word? The scene is always the same if the child conjectures the more inclusive interpretation (that is, if her first conjecture is animal rather than duck). For every time there is an observation that satisfies the conditions (whatever these are) for the appropriate use of duck, the conditions for the appropriate use of animal have been satisfied as well.

Analogous cases exist in the realm of verb meanings. To return to the instance dramatized by the blind learners, perceive, see, look, eye (in the sense of 'set eyes on'), face,

orient, pose the same subset problem. There is no seeing without looking, looking without facing, facing without orienting, etc. All this suggests that not only blind children, but sighted children as well, should have (essentially the same) difficulties in learning the meanings of look and see, because the distinction between the two words is not an observable property of the situations in which they are used. Yet, as I discussed earlier, it is just these "unobservable" properties that the blind and sighted three year olds held in common.

Gold (1967) addressed a problem that seems related to this one. He showed formally that learners who had to choose between two languages, one of which was a subset of the other, could receive no positive evidence that they had chosen wrong if they happened to conjecture the superset (larger) language. This is because the sentences they would hear, all drawn from the subset, are all members of the superset as well. It has therefore been proposed that learners always hypothesize the smaller (subset) language; they initially select the most restrictive value of a parameter on which languages vary (Berwick, 1981; Wexler and Manzini, 1987).

But the facts about the lexicon do not allow us to suppose that the child has a solution so simple as choosing the least inclusive possibility. In the end, they acquire all of them. Moreover, neither the most inclusive nor the least inclusive possibilities seem to be the initial conjectures; rather, some "middle" or "basic" level of interpretation is the one initially selected, i.e., duck and look (as opposed to mallard and glimpse seem to be the real first choices of the learners.[6]

In short, words that stand in a subset relation pose an intractable problem for an unaided observation-based learning procedure. This is because the child who first conjectures the more inclusive interpretation can receive no positive evidence from word-to-world mappings that can dissuade him. And the idea that he always begins with the least inclusive interpretation consistent with the data is falsified by the empirical facts.

Many semantic properties are closed to observation

But the verbs that most seriously challenge the semantic bootstrapping proposal still remain to be discussed: These are

---

6 These results can't be written off on grounds of the differential frequency of these words in the input corpus, for if the frequencies are changed the level of categorization does not. For instance, in some houses Fido is a more frequent word than dog, but in that case the youngest children think that the word meaning 'dog' is /faydo/ (Rescorla 1980).

the ones that don't refer to the observable world at all.

Locke noted that the meanings of many words involve properties that are closed to observation, but he did not consider this fact to be fatal to his overall position because his experience, partly warranted, was that those who used such "abstract" words didn't know what they were talking about half the time anyhow. Nevertheless a key problem for an unaided observational-learning story is that too many words that even a three or four-year old understands are related to the real world only in the most obscure and unobservable ways, if at all. Try, for example, to learn the meaning of the word think by titrating discourse situations into those in which thinking is going on, somewhere, when you hear /think/, vs those in which no thinking is happening. Remember that there isn't always brow-furrowing or a Rodin statue around to help. Keep in mind also that you are going to have to distinguish also among think, guess, wonder, know, hope, suppose and understand, not to speak of -- a few months or years later -- conjecture, figure, comprehend, discover, perceive, etc.

Many developmental psycholinguists rule such instances out of school on the grounds that these aren't words that children know very well at two and three years old, but this won't do. After all, we also want to understand the children who manage to survive to become the four and five year olds.

I don't really think this topic needs much more belaboring. If the child is to learn the meanings from perceptual discriminanda in the real world, the primitive vocabulary of infant perception has to be pretty narrow to bring the number and variety of data storing and manipulative procedures under control. But no such narrow vocabulary of perception could possibly select the thinkingness properties from events. I conclude that an unaided observation-based verb learning theory is untenable because it could not acquire think.

Summary

I've mentioned a number of problems for a theory that (solely or even primarily) performs a word-to-world mapping to solve the vocabulary learning task. These are that (i) such a theory fails to account for the fact that children whose exposure conditions are radically different acquire much the same representations of many words; (ii) plausible, narrowly drawn, candidates for event representation seem to be inadequate in accounting for the learning in certain apparently easy cases; (iii) broadening the hypothesis space so as to allow learners to distinguish among the many verb meanings may impose unrealistic storage, manipulation, and induction demands on the mere babes who must do the learning. In addition, (iv) many verbs are identical in all respects except the perspectives that they

adopt toward events or (v) the level of specificity at which they describe a single event; or (vi) don't refer to events and states that are observable at all. Since children learn the verb meanings despite these apparently formidible problems, my conjecture is that they have another source of information that redresses some of the insufficiencies of observation.

Part II: New approaches for vocabulary acquisition

How the blind child might have learned the visual terms

I return now to the problem Landau and I faced in understanding the blind child's semantic achievements. Keep in mind that the analysis of Table 1 was an attempt to explain only the most straightforward, perceptually relevant, aspect of her acquisition of look and see, namely that if these verbs had to do with haptic perception, there must have been pertinent objects close to her hands when her mother said those words. Yet even this simple idea seemed to be falsified by our analysis.

To find out why, our first step was to return to the data of Table 1 to see where and when the NEARNESS constraint had failed for so many uses of look and see. We found that the sentences that fell neatly under the object-nearby conjecture were very simple ones: If the mother had said something like

Look at this boot!
or See the apple?

invariably the boot or apple were NEAR, within the blind child's reach. But if the mother said

Let's see if Granny's home! (while dialing the phone)
Look what you're doing!
You look like a kangeroo in those overalls.
or Let's go see Poppy.

the "pertinent object" was likely to be FAR or there was NO such pertinent OBJECT intended. Clearly, the sentences that tripped up our simple story were queer ones indeed. The mother didn't seem in most of these cases to mean 'examine or apprehend' either haptically or visually, but rather 'determine', 'watch-out', or 'resemble.' Or else, as in the final example, a motion auxiliary (go) in the sentence transparently took off the NEARbyness requirement.

There are two ways to go now: One can claim that the NEARbyness environmental clue to the haptic interpretation was just a snare and delusion -- but that is ridiculous. It just HAS to be right that this aspect of the environment was part of what licensed the child's haptic interpretation. The other

choice is to find some non-question-begging way through which the child could have gotten rid of the sentences that otherwise would threaten the experiential conjecture. (The question-begging way, of course, is to say that the mother didn't mean 'haptically explore' in the offending sentences).

How can this be done? The clue is that not only the meaning, but the syntax too, of these offending sentences is special -- different from the syntax of sentences in which the child was <u>really</u> being told to explore and perceive nearby objects. This syntactic distinction may be available to the learner.

A syntactic partitioning of the verbs commonly used by the mother of the blind baby (based on the same corpus analyzed in Table 1), according to the subcategorization frames in which each verb appeared in the maternal corpus, is shown in Table 2; the verbs of Table 1 appear as the columns in this table, and the syntactic environments appear as the rows; the numbers in each cell are the number of instances of a verb in some particular syntactic environment.[7] Notice first that some of the typical syntactic environments for <u>look</u> and <u>see</u> are quite different from those for the other verbs in the set.

Moreover, we can -- with only a little fudging -- divide the environments of the vision-related verbs so as to pull apart those environments in which the NEARbyness contextual cue holds, and those in which it does not: That analysis is shown in Table 3. Essentially, the top rows of Table 3 show the maternal uses of these verbs in their canonical subcategorization frames (e.g., "Look at/see the frog," "Look up/down") and the deictic interjective uses that are the most frequent in that corpus (e.g., "Look!, That's a frog!" and "See?, That's a frog!"). When these syntactic types only are considered, the NEAR proportion of <u>look</u> rises (to 100%, from 73% in Table 1) and so does the NEAR proportion of <u>see</u> (to 72% from 39%). Thus if the learner can and does perform these analyses, the first result is that NEARbyness of the pertinent object becomes a much more reliable real-world clue than previously. But notice that the hypothesis now is that the child performs a sentence-to-world mapping, rather than the word-to-world mapping shown in Table 1: The child's interpretation of <u>extralinguistic</u> events has been significantly modulated by her attention to <u>linguistic</u> events, namely the syntax.

Landau and I made yet another, and much stronger, claim based on the kinds of outcomes shown in Table 2. This was that

---

[7] Specifically, the rows of this table represent <u>subcategorization frames</u>, the sister-nodes to V under the verb phrase.

	Group I		Group II				Group III			
	Look[a]	See[a]	Give[a]	Put[a]	Hold[a]	Play	Get[a]	Have	Go[a]	Come[a]
Look/see only										
Deictic										
V!	8									
V?		1								
V!, S	10									
V?, S		3								
V rel$_{here}$	2									
Other										
V like NP	5									
V S		5						1[b]		
come V NP		3								
Exclude *look/see*										
V NP PP			5	31	1		2			
V NP D				28	6		2			
V D NP				1						
V NP NP			16				2			
V rel$_{here}$				1						
Overlap with *look/see*										
V PP	3					7[c]			2	2
V D	2						5		10	13
V ∅	2	3							8	4
V NP		3			3	3	13	14		
V AP	2						3			
Totals	34	18	21	61	10	10	27	15	20	19

a. Verbs that occur with locative prepositions and adverbs.
b. A causative use of *have*: "Will we have Barbara come back, sit?"
c. *Play* with the nonlocative (reciprocal) preposition *with*: "You're not gonna play with the triangle, so forget it!"

Table 2: Subcategorization privileges of the common verbs used by the mother of a blind child during the learning period. The number in each cell represents the number of times that a verb is used in a particular frame environment (from Landau and Gleitman, 1985)

the range of subcategorization frames has considerable potential for partitioning the verb set semantically, and that language learners have the capacity and inclination to recruit this information source to redress the insufficiencies of raw observation. This examination of structure as a basis for deducing the meaning is the procedure we've called "syntactic bootstrapping." I turn now to a comparison of the hypothesis called "semantic bootstrapping" by Pinker to the one called "syntactic bootstrapping" by Landau and me.

The bootstrapping proposals compared

The two bootstrapping proposals are much alike in what they claim about correspondances between syntax and semantics, and are also alike in proposing that the child makes significant use of these correspondances. First I'll sketch, very briefly and informally, the kinds of syntactic/semantic correspondances that are crucially invoked in both proposals.

Syntactic/semantic linking rules: To an interesting degree, the structures in which verbs appear are projections from their meanings. To take a simple example, the different number of noun-phrases required by the verbs laugh, smack, and put in the sentences

(1) Arnold laughs.
(2) Arnold smacks Gloria.
(3) Gloria puts Arnold in his place.

is clearly no accident, but rather is semantically determined--by how many participant entities, locations, etc., the predicate implicates. Similarly, the structural positions of these noun-phrases relative to the verb also carries semantic information; thus, much more often than not the subject noun-phrase will represent the actor or causal agent (e.g., Arnold in sentence 1 and Gloria in sentence 2), and paths and goals will appear in prepositional phrases (in his place, in sentence 3). These links of syntactic position and marking to semantic properties, while by no means unexceptional, typify the ways that English represents semantic-relational structure. In short, verbs that are related in meaning share aspects of their clausal syntax. Zwicky (1971) put the idea this way:

"If you invent a verb, say greem, which refers to an act of communication by speech and describes the physical characteristics of the act (say a loud, hoarse, quality), then you know that...it will be possible to greem (i.e. to speak loudly and hoarsely), to greem for someone to get you a glass of water, to greem at your sister about the price of doughnuts, to greem "Ecch" at your enemies, to have your greem frighten the baby, to greem to me that my examples are absurd, and to give a greem when you see the explanation."

Semantic bootstrapping: Using the semantics to predict the syntax: As I mentioned earlier, both the bootstrapping proposals make critical use of these canonical relations between syntax and semantics. In the semantic bootstrapping procedure, the child fixes the meaning of a verb by observing its real-world contingencies. In Pinker's (1987) words:

> "...the child could learn verb meanings by (a) sampling, on each occasion in which a verb is used, a subset of the features...[8], (b) adding to the tentative definition for the verb its current value for that feature and (c) permanently discarding any feature value that is contradicted by a current situation."

I have argued at length that this position is too strong, for at least some features are unobservable. Yet no one can doubt that, at least sometimes, the context of use is so rich and restrictive as to make a certain conjecture about interpretation overwhelmingly likely.[9]

Once the verb meaning has been extracted from observation, the semantic bootstrapping hypothesis invokes the linking rules (the canonical syntactic/semantic mappings) to explain how the child discovers the structures which are licensed for the use of these verbs, much in the spirit of Zwicky's comments about the invented word greem. For instance, if a verb has been discovered to mean give, then it will appear in three-argument structures such as John gives the book to Mary. This is because the logic of 'give' implies one who gives, one who is given, and that which is given, and each of these entities requires a noun-

---

[8] The features are those mentioned in my earlier citation of Pinker (page 10 of this manuscript).

[9] At peril of making one argument too many, however, I can't resist complaining that Pinker's proposed procedure is too extreme. After all, sometimes the child is attending to one thing (say, the dog under the table) when the mother says something irrelevant to that (say, "Eat your peas, dear!"). So the learner better not "discard permanently" any feature that contradicts the current situation as he or she is conceiving it. In fact, positive imperatives pose one of the most devastating challenges to any scheme that makes word-to-world pairings for the mother will utter "Eat your peas!" if and only if the child is not then eating his peas. Thus a whole class of constructions seems to be reserved for saying things that mismatch the current situation.

phrase to express.

Not only is this position plausible. There is much evidence in its favor. Notably, Bowerman (1976; 1982) showed that children will make just such predictions about the syntactic structures licensed for verbs, presumably based on their prior fixing of the verb meanings: That evidence came from instances where children's conjectures were evidently too bold or insufficiently differentiated; that is, where they were wrong -- but still understandable. For instance, a subject of Bowerman's commanded "Don't eat the baby -- she's dirty!" on an occasion when the mother was about to feed the baby (whose diaper needed changing). Presumably, the child had conjectured that an intransitive motion verb (e.g., sink, as in The ship sank) could be uttered in a transitive structure (such as The captain sank the ship) to express the causal agent of this motion.

To summarize, the semantic bootstrapping procedure as developed by Grimshaw (1981), Pinker (1984) and others, works something like this: The child is conceived as listening to the words used, and then trying to figure out their meanings by observing their situational concomitants, the word-to-world pairing that I've discussed. Quoting Pinker (1984) again,

> If the child deduces the meanings of as yet uncomprehended input sentences from their contexts and from the meanings of their individual words, he or she would have to havve learned those word meanings beforehand. This could be accomplished by attending to single words used in isolation, to emphatically stressed single words, or to the single uncomprehended word in a sentence...and pairing it with a predicate corresponding to an entity or relation that is singled out ostensively, one that is salient in the discourse context, or one that appears to be expressed in the speech act for which there is no known word in the sentence expressing it (p. 30).

Once the meanings have been derived from observation, the child can project the structures from her (innate) knowledge of the rules that map semantic structures onto syntactic structures (by procedures variously called mapping rules,linking rules, projection rules, or semantic redundancy rules). Perhaps so, but I have been arguing that entities and relations cannot in general be singled out ostensively, that "salience" and the question of what's "expressed in the speech act" are not so easily recoverable as this perspective must insist. For such reasons, Landau and I developed a procedure that looks quite different from this.

Syntactic bootstrapping: The syntactic bootstrapping proposal in essence turns semantic bootstrapping on its head. According to this hypothesis, the child who understands the

mapping rules for semantics onto syntax can use the observed syntactic structures as evidence for deducing the meanings. The child is conceived as having certain concepts in mind, say, 'look' or 'put', and is engaged in a search for the words that express these concepts. To accomplish these aims, the child observes the real-world situation but also observes the structures in which various words appear in the speech of the caretakers. That is to say, the child performs a sentence-to-world pairing rather than a word-to-world mapping. Such a procedure can succeed because, if the syntactic structures are truly correlated with the meanings, the range of structures will be informative for deducing which word (qua phonological object) goes with which concept. Such a procedure will be quite handy if, as I have argued, raw word-to-world mapping cannot succeed. The difference between semantic bootstrapping and syntactic bootstrapping, then, is that the former procedure deduces the structures from the word meanings that are antecedently acquired from real-world observation; while the latter procedure deduces the word meanings from the semantically relevant syntactic structures associated with a verb in input utterances.

Let us take the simple examples of put, look, and see., which occurred in the corpus provided by the blind child's mother. Verbs that describe externally caused transfer or change of possessor of an object from place to place (or from person to person) fit naturally into sentences with three noun-phrases, e.g. John put the ball on the table. This is just the kind of transparent syntax/semantic relation that every known language seems to embody and therefore may not be too wild to conjecture as part of the original presuppositional structure that children bring into the language learning task (Jackendoff, 1978; 1983; Talmy, 1975; Pinker, 1984). That is, 'putting' logically implies one who puts, a thing put, and a place into which it is put; a noun-phrase is assigned to each of the participants in such an event. In contrast, since one can't move objects from place to place by the perceptual act of looking at them, the occasion for using look in such a structure hardly, if ever, arises (John looked the ball on the table sounds unnatural). Hence the chances that /put/ means 'put' are raised and the chances that /put/ means 'look' are lowered by the fact that the former and not the latter verb appears in three-noun-phrase constructions in caretaker speech (see Table 2).[10]

---

[10] The exceptions are (1) if you believe in psychokinesis or (2) if the rules of some game make it so that, in effect, an external agent can cause an object to move by looking at it, e.g., The shortstop looked the runner back to second base. In effect, once look does mean cause-to-move-by-perceptually-exploring, it becomes comfortable in this construction. Of course these simple examples vastly underestimate the detail required if such a theory is to become viable. One such

	Near	Far	No object	"Near" proportion
Canonical sentence frames and deictic uses				
Look at NP	3	0	0	
Look D	2	0	0	1.00
Look!	8	0	0	
Look! this is NP	10	0	0	
See NP	1	2	0	
See?	1	0	0	.72
See?, This is NP	3	0	0	
With motion auxiliaries				
Come see NP	0	3	0	.00
Other environments				
Look AP	0	1	1	
Look like NP	0	0	5	.18
Look how_{wh}	0	2	0	
Look ∅	2	0	0	
See S	2	3	0	.25
See ∅	0	2	1	
Total (all environments)				
Look	25	3	6	.73
See	7	10	1	.39

Table 3: Situational contexts for the common verbs used by the blind child's mother, organized according to the syntactic (subcategorization frame) contexts (from Landau and Gleitman, 1985)

Verbs of perception and cognition are associated with some other constructions, as they should be. For example, if a verb is to mean 'see' (perceive perceptually), it should appear with noun-phrase objects as in John saw a mouse, for noun-phrases are the categories that languages select to describe such entities as mice. But since events as well as entities can be perceived, this verb should also appear with sentence complements, since clauses are the categories selected by languages for expressing whole events (e.g., Let's see if there's cheese in the refrigerator). The possibility that /see/ means 'see' is increased by appearance in this construction, and the likelihood that /put/ means 'see' is decreased by the fact that one hardly, if ever, hears Let's put if there's cheese in the refrigerator; see again Table 2).

Speaking more generally, certain abstract semantic elements such as 'cause,' 'transfer,' and 'cognition' are carried on clause structures (subcategorization frames) rather than (or in addition to) as item-specific information in the lexical entries of verbs. These semantically relevant clause structures will be chosen for utterance only to the extent that they fit with the overall meanings of the verb items. It follows that the subcategorization frames, if their semantic values are known, can convey important semantic information to the verb learner. To be sure, the number of such clause structures is quite small compared to the number of possible verb meanings: It is reasonable to assume that only a limited number of highly general semantic categories and functions are exhibited in the organization that yields the subcategorization frame distinctions. But each verb is associated with several of these structures. Each such structure narrows down the choice of interpretations for the verb. Thus these limited parameters of structural variation, operating jointly, can predict the possible meaning of an individual verb quite closely. Landau and Gleitman showed that the child's situational and syntactic input, as represented in Tables 2 and 3, were sufficient in principle to distinguish among all the verbs commonly used in the maternal sample for the blind child. This general outcome is schematized in Figure 3.

The potential virtues of this syntactically informed verb-

---

problem is that the child must impose the proper parse on the sentence heard, lest John saw the book on the table be taken as a counter-example (that is, the analysis is to be of sister-nodes under VP only, and a theory of how the child determines such configurations antecedently is a requirement of the position). Another real difficulty is that the child might run into one of many quirky constructions like John saw his brother out of the room, looked his uncle in the eye, etc.

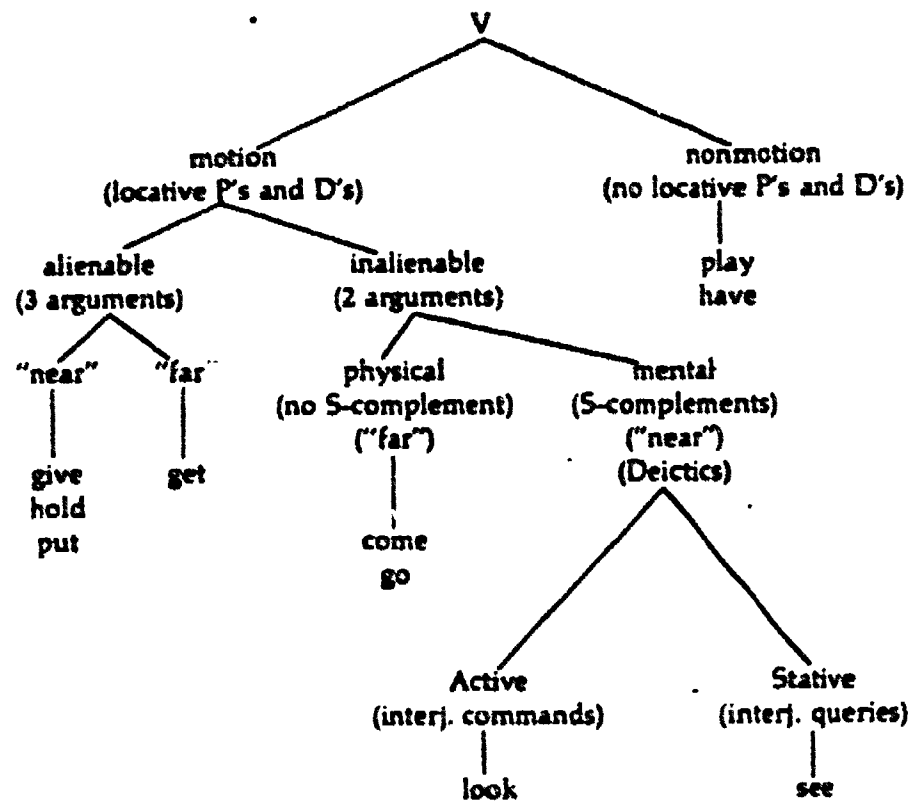

Figure 3: Summary of the situational and syntactic distinctions among verbs commonly used by the mother to the blind child during the learning period. (from Landau and Gleitman, 1985)

learning procedure are considerable. First, it serves the local purpose of offering a non-magical explanation for the blind child's acquisition of visual terms, as just described. Second, it points the way toward acquisition of terms when observation fails. This is because, for example, mental verbs such as <u>think</u> are unambiguously marked by the syntax (by taking sentence complements) even though their instances cannot be readily observed in the world. Third, it gives the child a way of learning from a very small database. This is because the number of subcategorization frames associated with each verb is small (on the order of 10 - 20), and these are the data requirements for the procedure to work. Fourth, that database is categorical rather than probabalistic: Though verb uses to the child are often pertinent to what is going on in the here-and-now, sometimes they are not (e.g., the mother may speak of running to the store while she sits in her parlor). In contrast, mothers virtually never speak ungrammatically to their children -- that is, use verbs in nonlicensed syntactic environments (Newport, 1977). Thus the child can take one or two instances of a verb in some frame as conclusive evidence that it is licensed in this environment. Finally, what is used in this procedure for learning is part of what must be known by an accomplished speaker: Knowing the subcategorization privileges for each verb is part of what it means to be an English speaker. In contrast, many of the situational analyses constructed along the way by the semantic bootstrapper will not figure in the final definition of a verb.

In the light of all these virtues, it would be nice if this theory turned out to be part of the truth about how the verb vocabulary is acquired. I will provide some empirical evidence in its favor below. But first some presuppositions of the position have to be defended before so apparently "abstract" a procedure can be considered viable at all. I turn now to such questions.

<u>Prolegomena to the bootstrapping hypotheses</u>

The bootstrapping hypotheses involve a number of presuppositions that require demonstration in their own right, lest all learning questions be begged. In company with all known theories of word learning, they presuppose that the human child, by natural disposition (or learning during the prelinguistic period) is able to conceive of such notions as 'running' and 'looking' and implicitly understands that words make reference to such acts and events. Past this background supposition, both semantic and syntactic bootstrapping procedures -- but especially the latter -- make very strong claims about the child's knowledge <u>as verb learning begins</u>. I will now go through these claims, mentioning some of the experimental evidence that gives them plausibility.

Are the rules linking semantics and syntax strong and stable enough to support a learning procedure? If the syntactic structures associated with verbs are uncorrelated with -- or hardly correlated with -- their meanings, then the child can't learn much about the meanings by observing the structures. No one doubts the sheer existence of such form/meaning regularities owing to the results achieved by a generation of linguists, notably Gruber, Fillmore, Vendler, Jackendoff, and Levin (and many others), but questions can be raised about the stability, degree, and scope of these relations. That is, how far can a syntactic analysis such as that in Table 2 succeed in partitioning the lexicon semantically for the child learner?

I'll mention one line of investigation of these questions from our laboratory. Fisher, Gleitman, and Gleitman (in press) reasoned as follows: If similarity in the range of subcategorization frames of verbs is correlated with similarities in their meanings, then subjects asked to partition a set of verbs (a) according to their meanings and (b) according to their licensed structures should partition the verbs in much the same ways. To test this idea, one group of subjects made judgments of meaning-similarity for triads of verbs presented to them. Specifically, they chose the semantic outlier in each triad (e.g., shown eat, drink, and sing, they choose sing as the outlier, but shown eat, drink, and quaff they might choose eat). A semantic space for a set of verbs was derived from these data by tabulating how often two verbs stayed together (were not chosen as outlier) in the context of all other verbs with which they were compared. Presumably, the more often they stayed together, the more semantically similar they were. A second group of subjects gave judgments of grammaticality for all these same verbs in a large number of subcategorization frames. A syntactic space was derived in terms of the frame overlap among them. The similarity in the syntactic and semantic spaces provided by these two groups of subjects was then compared statistically.

The finding was that the frame overlap among the verbs is a very powerful predictor of the semantic partitioning. In short, verbs that behaved alike syntactically were, to a very interesting degree, the verbs that behaved alike semantically. Such results begin to show that a syntactic partitioning of the input can provide important evidence for a learner who is disposed to use such information -- as was conjectured for the blind child, see Figure 3.

Are the semantic/syntactic relations the same cross-linguistically? The first proviso to the conclusion just drawn, for learning questions, is that the semantic-syntactic relations have to be about the same across languages. Otherwise, depending on the exposure language, different children would have to perform different syntactic analyses to derive aspects of the meaning. And that, surely, begs the questions at issue.

Recent theorizing in linguistics does support the idea that there are semantic/syntactic linkages that hold across languages. In a recent version of generative grammar (Government/-Binding theory; see Chomsky, 1981), some of these relationships are stated as universal principles of language design. One example is the mapping of entities implied by the verb logic one-to-one onto noun-phrase positions in the clause: Every NP in a sentence must receive one and only one thematic role (the theta-criteri- on). Moreover, a related principle (the projection principle) states that the theta-criterion will hold at every level of a derivation; in particular, that argument structure is preserved on the surface clause structures. This is just the organization required by a bootstrapper -- semantic or syntactic.

Talmy (1975; 1985) has investigated a number of typologically quite different languages and found a variety of striking similarities in how their semantics maps onto the syntax. For those who prefer experimental evidence from linguistically naive subjects, Fisher et al, in a very preliminary cross-linguistic foray with their method, showed that the relationship between being a verb of cognition and accepting sentence complements is as strong and stable in Italian as in English.

The two relationships just mentioned (that a NP is assigned to each participant in the event, and that verbs encoding the relation between an agent and a propositon accept sentence complements) are not only true cross-linguistically. They have a kind of cognitive transparency that makes them plausible as part of the presuppositional structure children might really bring into the language learning situation. As Jackendoff puts this point:

> In order to lighten the language learner's load further, it seems promising to seek a theory of semantics (that is, of conceptualization) in which the projection rules are relatively simple, for then the child can draw relatively straightforward connections between the language he hears and his conception of the world. The methodological assumptions for such a theory would be that syntactic simplicity ideally corresponds to conceptual simplicity; grammatical parallelisms may be cluses to perceptual parallelisms; apparent grammatical constraints may reflect conceptual constraints.
>
> (1978; p. 203)

From these and related arguments and demonstrations, I think the plausibilty of the bootstrapping theories receives at least some initial defense.

Can the learner analyze the sound wave in a way that will

support discovery of syntactic structure? There is a timing difference in the requirements of semantic and syntactic bootstrapping approaches: For the latter approach, the learner has to be able to parse the sentences that she hears in order to derive a syntactic analysis; moreover, at least some of the mapping rules have to be in place before the verb meanings are known and thus the whole game is over. There is strong evidence supporting both these claims:

Can infants parse?: Once upon a time, not so very long ago, it was believed that babies could divide up the sound wave into words but not into phrases. This perspective necessitated complex theories for how learners could derive phrasal categories from the initial word-like representations (Wexler and Culicover, 1980; Pinker, 1984). In retrospect, these ideas were somewhat improbable. For one thing, there is evidence that infants are sensitive to such physical properties of the wave form as change in fundamental frequency, silent intervals, and syllabic length, all of which are universal markers of phrase boundaries (see, e.g., Fernald, 1984). As Gleitman and Wanner (1982) pointed out, the physical correlates of word segmentation are far more subtle and less reliable. More generally, our reading of the cross-linguistic facts about language learning led us to propose that the infant's analysis of the wave form was as a rudimentary phrase-structure tree.[11] In a similar vein, Morgan and Newport (1981; Morgan, Meier, and Newport, 1988, showed in a series of artificial language-learning experiments that adults could learn phrase structure grammars if provided with phrase-bracketing information but not if provided only with word-level information. This finding led these investigators independently to the same proposal about the child's initial representation of the input wave forms. Recently, Hirsh-Pasek and her colleagues (1988a) have shown that prelinguistic infants listen to maternal speech doctored so as to preserve phrase- and clause-bounding information in preference to speech doctored so as to remove or becloud this information (see Gleitman et al, 1987, for a review of the evidence and its interpretation for a language acquisition theory).

---

[11] Notoriously, word-segmentation in a language like English is so fraught with ambiguity that new pronunciations (e.g., nother and apron replacing other and napron) are quite common. Moreover, there are long-lasting errors by children, e.g., one six-year old wrote "The teacher said, Class be smissed!" The phrasal parses suggested by Gleitman and Wanner were "rudimentary" to the extent that the unstressed elements in the phrases were presumed to be less well analyzed than the stressed elements, and the phrases were unlabelled (but see Joshi and Levy, 1982, for evidence that much of labelling, or its equivalent, can be derived from "skeletal" representations in which there are configurations but no overt labels).

The evidence just cited is not precise enough to give a detailed picture of the infant's phrasal parse[12] However, that evidence is strong enough to support the view that children, even in the prelinguistic period, impose an analysis on the wave form sufficient for partitioning it into phrases. There is weaker but still suggestive evidence that the young learners also have the wherewithal to label the phrases differentially (see again footnote 11). It is incontrovertible that the two and three year olds who are the real verb learners can achieve the analyses of input shown in Table 2, and which are a requirement for achieving the semantic partitioning of the verb set shown in Figure 3.

Does the learner know the syntactic/semantic correspondance rules? A crucial further requirement for the bootstrapping hypotheses is that the child understand the semantic values of the subcategorization frames. A child who recovers the meaning from observation, and who is to deduce the structures therefrom, has to know what the semantics of the verb implies about the syntactic structures licensed. And a child who recovers the syntactic structures licensed for verbs from the linguistic contexts in which she hears them has to know what semantic elements are implied by participation in these structures. As Jackendoff emphasized, the burden of learning would certainly be reduced for a child in possession of such information. But do real learners actually have it? There is striking evidence that they do.

Golinkoff et al (1987) developed a very useful paradigm for studying very young children's comprehension. Essentially, they adapted a procedure designed by Spelke for studying infant perception. The set-up for the language case is shown in Figure 4. The child sees different scenes displayed on two video screens, one to the child's left, one to her right. The scenes are accompanied by some speech stimulus. The mother wears a visor so that she cannot observe the videos and so cannot give hints to her child. Hidden observers are so positioned that they cannot observe the video, but they can observe which way the child is looking, and for how long. It turns out that children look sooner and longer at the video that matches the speech input.

In a first demonstration relevant to the syntactic bootstrapping hypothesis, Golinkoff et al showed that 19-month old children -- many of whom had never put two words together in an utterance, and knew few if any verbs -- understand some facts about the semantic values of English constructions. Two

---

[12] But see Eccles and Newport, forthcoming, for experimental findings that support significant theorizing in this area.

simultaneous videos showed cartoon characters known to the children interacting. For some subjects, the stimulus sentence was Big Bird is tickling Cookie Monster. For the others, it was Cookie Monster is tickling Big Bird. The children demonstrated by their selective looking that they knew which sentence described which observed event: They looked longer at the screen showing Big Bird tickling Cookie Monster when they heard the former sentence, and at the screen showing Cookie Monster tickling Big Bird when they heard the latter sentence. That is, these children recognize the order of phrases (or something approximating phrases) within the heard sentence and also understand the semantic significance of the ordering for the propositional interpretation of English speech (see also Slobin and Bever, 1982, for cross-linguistic evidence on this topic).

I and my colleagues (Hirsh-Pasek et. al, 1988b) used this same procedure to investigate one more property of the mapping rules, namely the causative structure for which Bowerman (1982) had found many innovative uses by youngsters: Roughly, intransitive motion verbs (e.g., Big Bird turns) can be "transitivized" in English and then will express the causal agent as well (Cookie Monster turns Big Bird).

To study this question using the procedure of selective looking, it is necessary that both entities appear in the stimulus sentence; otherwise the children may use the relatively trivial strategy of looking at the stimulus showing Big Bird if and only if Big Bird is mentioned. Hence the real stimuli used were, for example, Big Bird is turning Cookie Monster and Big Bird is turning with Cookie Monster. One video showed the two characters turning side by side, and the other video showed one character physically causing the other to turn. In addition to verbs like turn that (by maternal report) were probably known to the 2-year old subjects, unknown ones were also used. For example, the characters were shown flexing their arms, or one flexing the arms of the other, along with the stimuli Big Bird is gorping Cookie Monster and Big Bird is gorping with Cookie Monster. We were unable to show stable effects of the syntactic structure for children at 24 months of age. But just about every youngster by 27 months showed the effect of the structure, by looking longest at the syntactically congruent screen.

The conclusions to be drawn are very important ones for the syntactic bootstrapping hypothesis. The paired actions are the same, e.g., both are of turning in a circle, or both are of flexing the arms. What differs is whether a causal agent of that action is also present in that scene. The children seem to know that only the transitive use of the verb can be expressing that cause. More strongly, that causal agent cannot be in an oblique argument position (the with phrase).

Prior demonstrations of knowledge of mapping rules have generally been with much older children. For instance, Bowerman notes that most spontaneous overgeneralizations of the causative structure ("Don't eat the baby!") are later, in the three to five-year old period.[13] Pinker and his colleagues have offered many compelling demonstrations of a variety of mapping rules, but again mainly with three to five year olds. These findings give general support to the idea that learners recruit the semantic/syntactic correlations somewhere during the course of learning. But the early appearance of these skills is crucial as support for the notion that the child has the mapping rules under control early enough for them to contribute to the acquisition of the verb meanings themselves.

Using syntax to acquire verb meanings

So far I've tried to show that a number of presuppositions of syntactic bootstrapping are reasonable: The language does exhibit strong and stable syntactic/semantic correlations, and these powerfully predict adult classificatory behavior; children in the prelinguistic period can and do parse sentences to recover the analyses required for extracting subcategorization frame information; such phrasal information is a requirement for language learning, at least for adults in the artificial language-learning laboratory; children at a very young age and language-learning stage understand the semantic values of at least some syntactic frames.

All of these findings were prolegomena to the syntactic bootstrapping approach. They were adduced because it is bad enough that this approach seems so unnatural and formal a one for a child to choose; it would be worse if the child couldn't come up with the analyses that the position presupposes. But now that I've presented at least some preliminary support that children can meet these prior requirements, the question remains: Do they use syntactic evidence to decide on the meaning of a new word?

---

13 But see also Naigles, Gleitman, and Gleitman (in press) for a demonstration that two year olds understand the significance of new motion transitives, even though they may not be brave enough to invent any until they are three. The subjects here were asked to "act out" scenes using a Noah's Ark and its animal inhabitants. For instance, the child might be told to act out "Noah brings the elephant to the ark." But some of the stimuli were more unusual, e.g., "Noah comes the elephant to the ark" or "The elephant brings to the ark." The children by their acting-out performances showed that they thought transitive come means 'bring' and that intransitive bring means come.

The first, and justly famous, work on this topic was done by Roger Brown (1957). He showed three to five year olds a picture in which, say, spaghetti-like stuff was being poured into a vessel. Some subjects were asked to show some gorp, others a gorp, and still others gorping. The subjects' choices were, respectively, the spaghetti, the vessel, and the action. Evidently, the semantic core of the word classes affects the conjecture about the aspect of the scene in view that is being labelled linguistically.

Brown's result, though alluded to respectfully, just sat there for twenty years or so because in this respect as in many others Brown was a theorist ahead of his time. Eventually, MacNamara took up and advanced these ideas: In his important 1972 paper, he argued forcefully for the place of language structure in language acquisition. Experimentally, Baker, Katz, and MacNamara (1974) showed that children as young as 19 months used the structure in which new nouns appeared (a gorp vs Gorp) to decide whether a new word encoded a class or an individual (i.e., a doll of the gorpy type, or some doll named Gorp). Thus the lexical category assignments of words were shown to carry semantic implications, and these were evidently recruited by learners.

Naigles (in press), working in my lab and also in the lab of Hirsh-Pasek and Golinkoff at Temple University, extended this kind of demonstration to the case of verb learning (that is, to the usefulness of syntax for drawing semantic inferences within a single lexical category), thus giving the first direct demonstration of syntactic bootstrapping at work.

Twenty-four month olds were again put into the selective looking situation. This time, however, their task was to decide between two utterly disjoint interpretations of a new verb. In the training (learning) period, they saw a single screen, and the following mad event: A rabbit is pushing a duck down into a squatting position with his left arm (these were people dressed up as rabbits and ducks so they did have arms). The duck pops up, and the rabbit pushes him down again, etc. Simultanously, both rabbit and duck are making big circles in the air with their right arms. Some children heard a voice say The rabbit is gorping the duck and other children heard The rabbit and the duck are gorping as they watched this scene.

Succeeding the observation, the screen goes dark and the voice is heard to say something syntactically uninformative, e.g. Oh, there's gorping; now there's gorping! Now new videos appear on two screens, as shown in Figure 4. On one screen, the rabbit is pushing the duck down (but with no arm-wheeling). On the other screen, rabbit and duck are wheeling their arms (but with no squatting or forcing to squat). The child's looking time at the screens, as a function of his syntactic

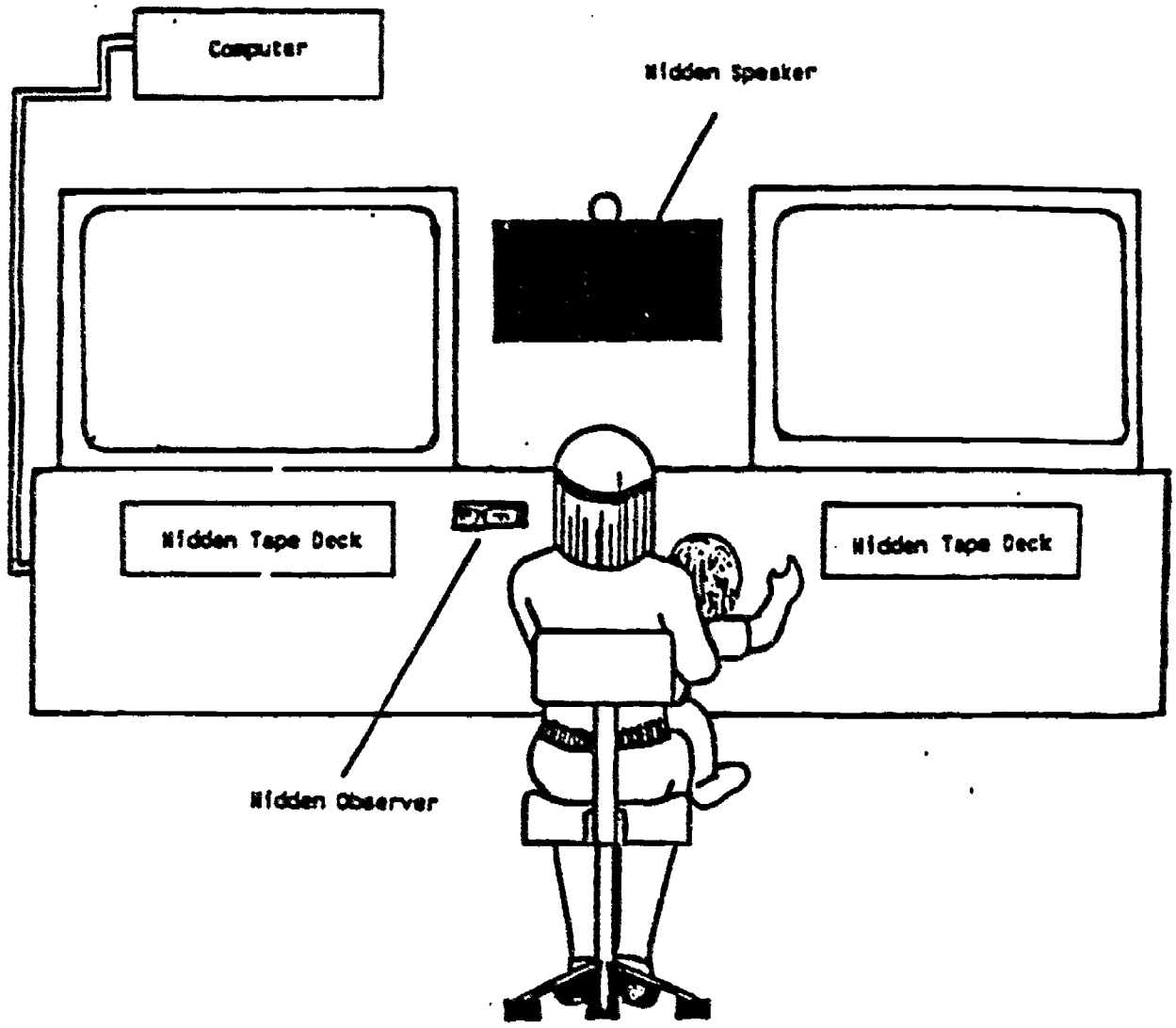

Figure 4: Set-up for the selective looking experiments (from Naigles, in press)

introducing circumstances, is now recorded (double-blind as usual, i.e., neither the mother nor the experimenters know which screen the child saw during the training period).

Naigles' result was that virtually every 24-month old tested -- and there were many, this being a Ph.D. thesis-- showed the effect of the syntactic introducing circumstance. Those who heard the transitive sentence apparently concluded that gorp means 'force-to-squat.' Those who heard the intransitive sentence concluded that gorp means 'wheel the arms.'[14]

What shall we conclude from this experiment? Clearly the child uses the event-context in some way to license conjectures about a verb meaning. But in this case, "The Main Event" is ambiguous not only in principle but in fact. Under these trying circumstances, at least, the learner attends to the information potential of the semantically relevant syntactic evidence.

A question of scope

So far the experiments I've mentioned have lingered nernervously around a few constructions, e.g. the lexical causative in English which is a notorious focus of syntactic extension by adults as well as children. Even if it is accepted that children sometimes do use syntactic evidence to bolster their semantic conjectures, how broad can the scope of such a procedure be? Maybe its role is just to clean up a few little details that are hard to gleen from the world -- just backwards semantic bootstrapping, as Pinker has sometimes put the matter.

The relative roles of linguistic and extralinguistic observation as the source of word-meaning acquisition is not within calling distance of settlement, of course. But the

---

[14] Notice that in all the selective looking experiments I've mentioned all the participants are animate so there's no room for counter-interpretations such as the strategy of assigning the animate entity to the subject position. Note also that in the present experiment the intransitive sentence contained a conjoined nominal (The duck and the rabbit) and this might be seen as a defect: Maybe the child knows the difference beteen a preverbal and a postverbal nominal rather than the difference between a transitive and an intransitive structure. This interpretation is effectively excluded by the version presented earlier (Hirsh-Pasek et al, 1988b) in which the two noun-phrases appear in different argument positions, one serially before and one after the verb (Big Bird is turning with Cookie Monster). For elegance, however, it certainly would be nice to redo the present experiment with the stimulus type used in the former one.

burgeoning linguistic and psycholinguistic literature on lexical semantics suggests that the semantic/syntactic linkages may be quite pervasive and stable, and play a potent role in organizing the verb lexicon.

Fisher, Hall, Rakowitz, and I have just completed some studies designed to investigate the scope of children's exploitation of the syntactic environment in learning new verb meanings. I believe that our prior studies with children two years old and younger yield evidence that satisfies an explanatory demand of this approach: The bootstrapping procedure has to be able to operate very early in the child's linguistic life, else it hardly explains how verbs are acquired.

Nevertheless, the selective looking paradigm (which is one of very few that work with toddlers) is too much of a straightjacket to be the only vehicle for extensive investigation of this approach. It is tedicus in the extreme to set up (requiring the preparation of movies, etc.), takes hoards of infants to carry out (for some scream or sleep or worse and have to be removed from the premises; and only a few trials can be presented even to the more docile infants), and yields probabilistic results (in part because the subjects are not notified directly of the task they are to perform). Moreover, it may very well be that the child's knowledge of the linking rules expands as his language knowledge grows, creating more latitude within which he can learn new meanings from linguistic evidence (After all, in the end we can do it by looking in the dictionary).

We therefore set out to see whether preschoolers (aged 3 and 4 in the version now presented) would give us meanings in response to linguistic/situational stimuli upon request. The idea derived from a manipulation attempted by Marantz (1982). He had asked whether children are as quick to learn noncanonical vs noncanonical mappings of semantics onto syntax. He introduced children to novel verbs as they watched a movie. For instance, one movie showed a man pounding on a book with his elbow. Marantz' question was whether children were as quick to learn that <u>The book is moaking Larry</u> (the noncanonical mapping) was a way of describing this scene as that <u>Larry is moaking the book</u> (the canonical mapping) was a way of describing the scene.

Although the manipulation was an interesting one, unfortunately Marantz never asked the children how they interpreted the scene, so his results are not really relevant to understanding the child's perception of syntactic/semantic correlations. That is, Maranz <u>presupposed</u> that a scene viewed has only a single interpretation, an idea I have strenuously opposed throughout this discussion. My colleagues and I revised this experiment, changing the measure so we could find out about the child's comprehension in these circumstances. In essence, we

asked how the nonsense word is interpreted within differing linguistic environments. As a first step, we showed the moaking scene (in which Larry pounds the ball with his elbow) to adults. If we said "This scene can be described as a "moaking scene" and then asked them what moak meant, they said "pounding." And if instead we showed them the scene and said "This is Larry moaking the book," they still asserted that moak means "pound." But when we showed them the scene and said "This is the book moaking Larry," they answered that moak means "hurt" or "resist."

This suggests that adults make use of the fact that particular surface syntactic structures are associated with particular semantic values. They seem to bootstrap the meaning from examination of the scene taken together with its syntactic expression, just as the syntactic bootstrapping procedure claims. To be sure, the contextless presentation of moak with this scene irresistably yields the concept 'pound' as its interpretation. So there's much to be said for the idea of "salience" in the interpretation of events (though, to be sure, no one knows what exactly). But the important point is that there is a categorical shift in interpretation of the same scene -- to a less salient, but still possible, interpretation -- in response to its linguistic setting; namely 'pound' if Larry is in the subject position, but 'hurt' if the book is in that position.

Fisher et al now adapted this procedure for children. We took advantage of the idea, popularized by such Penn developmentalists as Gelman, Waxman, Macario, and Massey, that preschoolers will do just about anything to help out a puppet. We introduced a puppet saying "This puppet sometimes talks puppet-talk so I can't understand him; can you help figure out what he means?" The children were happy to oblige. They were shown videotaped scenes in which animals were performing certain acts. For example, a rabbit appeared, looked to the left, and then ran rapidly off the screen toward the right; directly behind him ran a skunk, also disappearing at the right. Then the child would hear either "The rabbit is gorping the skunk" or else "The skunk is gorping the rabbit."

The structures investigated are shown in Table 4. They are designed to ask whether the child is sensitive to the number of argument positions (stimuli 1 and 2), the structural positions of agent and patient (stimuli 3 and 4), and the structural positions taken together with prepositional markers of the oblique roles (stimuli 5 and 6). Thus we now began to investigate the scope of the structural/semantic linkages to which learners may be sensitive. Notice that the pairs chosen are just the kind that I have discussed throughout: The same scenes, multiply interpretable, are shown but accompanied by a novel verb used in varying constructions.

SCENE	STIMULUS SENTENCE
1. a) Rabbit eating.	The rabbit moaks.
b) Elephant feeding rabbit	The elephant moaks the rabbit.
2. a) Monkey pushing elephant.	The monkey pumes the elephant.
b) Elephant falling	The elephant pumes.
3. a) Monkey riding elephant.	The monkey gorms the elephant.
b) Elephant carrying monkey.	The elephant gorms the monkey.
4. a) Rabbit fleeing skunk.	The rabbit zarps the skunk.
b) Skunk chasing rabbit.	The skunk zarps the rabbit.
5. a) Rabbit giving a ball to elephant.	The rabbit ziffs a ball to the elephant.
b) Elephant taking a ball from rabbit.	The elephant ziffs a ball from the rabbit.
6. a) Skunk putting blanket on monkey.	The skunk is biffing a blanket on the monkey.
b) Skunk covering monkey with a blanket	The skunk is biffing the the monkey with a blanket.

Table 4: Stimuli used by Fisher, Hall, Rakowitz, and Gleitman (forthcoming). All Ss were exposed to the same six scenes (each scene has two plausible interpretations, called a) and b) in the left-hand column. Along with these scenes, half the children heard a) stimulus sentences and half heard b) stimulus sentences (with appropriate counterbalancing across Ss and stimuli).

The findings are shown in Table 5 They are presented in terms of the likelihood of various responses depending on the introducing syntactic structure. For example, the response give (response A) to structure (a) in Table 4 (The rabbit ziffs a ball to the elephant) was made by 4 Ss, but the response take (or, equivalently, get) was made by only 2 Ss in this condition. Symmetrically, the response take or get (response B) was made by 5 Ss in response to structure (b) in Table 4 (The elephant ziffs a ball from the rabbit), while that response was never made to structure (a).

Overall, 71 relevant responses made by these children were congruent with the semantic value implied by the syntactic structure, while only 13 relevant responses were inconsistent with the structural information. Moreover, for each scene and for each syntactic type, the number of syntactically congruent responses is greater than the noncongruent responses. The level of congruence was about the same for all three semantic/syntactic relations studied: 83% congruent responses when the variable was number of noun-phrases, 89% congruent responses when the variable was structural position of these noun-phrases, and 81% congruent responses when the variable was position plus prepositional marking.

One might object that these children are "merely" paraphrasing verbs that they previously know to occur in these syntactic environments. That is true, but it does not take away seriously from our interpretation of these findings: The children knew, evidently, that the appropriate paraphrase had to be one which fit both with the scene and with the sentence structure heard. This is the reverse of Pinker's claim that the verb meanings must be acquired by extralinguistic observation in advance of, and as the basis for, deducing their appropriate syntactic structures. But the results are exactly those expected in the syntactic bootstrapping approach.

A note on the input corpus

One of several holes in our present evidence has to do with the characteristics of caretaker speech. I have presented a single example corpus (Table 2) tending to support the idea that caretaker speech is rich enough to yield quite a full range of structures to support the syntactic bootstrapping procedure. And this corpus was for a mother speaking to a blind child, whose word-learning situation may be quite special. We are now analyzing an extensive corpus of mother/child speech in a naturalistic setting (originally collected by Landau and Gleitman) to see whether children characteristically receive the range of structures adequate to support a realistic syntax-based procedure (Lederer, Gleitman, and Gleitman, 1989). So far, the prospects from this larger database look good. Lederer finds

Syntactic type of the stimulus		Response A		Response B	
a	b	S_a	S_b	Sb	S_a
eat	feed	7	1	6	1
push	fall	8	3	4	0
ride	carry	7	2	4	0
flee	chase	6	0	8	1
give	take	4	2	5	0
put	cover	8	3	4	0
TOTALS:		40	11	31	2

Table 5: 16 Ss (aged 3-4) asked: WHAT DOES BIFFING MEAN? Not all subjects answered every question, accounting for totals in each row not totalling to 16. Also, some responses were irrelevant to either interpretation of a stimulus, e.g., S might say in response to the flee/chase scene "They're having fun!" These irrelevant stimuli are excluded from this tabulation.

that each of the 24 verbs most often used by these mothers to their children has a distinctive syntactic distribution. When the usages are pooled across mothers, these distinctions are preserved.

The next question is whether these syntactic distributions map onto a semantic space coherently. An independent assessment of the semantic relations among these verbs is required as the evidence. Lederer therefore is now testing this issue by using these verbs in the kind of manipulation employed by Gleitman et al; namely, asking adult subjects for judgments of the semantic outlier in all triads of these verbs. Preliminary inspection of the verbs suggests that the semantic clusters that emerge from these data are strongly predicted by the syntactic overlaps in the maternal corpora.

Conclusions

I began discussion by acknowledging the intuitive power of Locke's view that words are learned by noticing the real-world contingencies for their use. Then I tried to show that such a word-to-world mapping, unaided, was in principle insufficiently constrained to answer to the question of how the child matches the verbs (qua phonological objects) with their meanings. The solution that I and my colleagues have offered was that semantically relevant information in the syntactic structures could rescue observational learning from the sundry experiential pitfalls that threaten it. This theory, of course, is the very opposite of intuitive. But when probable solutions fail, less probable ones deserve to be considered. I therefore sketched a rather wide-ranging empirical review that we have undertaken to see whether, after all, children might not be deducing some of the meanings from their knowledge of structural/semantic relations. I believe that the evidence we now have in hand materially strengthens the plausibility of the viewpoint.

Still, the conclusions that can be drawn currently about the generality and pervasiveness of syntactic bootstrapping must be exceedingly tentative, on a variety of grounds. Some of these I have discussed: No one has more than a glimmer of an idea about just how the verb lexicon is organized, and therefore we don't really know how much information about semantics can be gleaned from that organization. Also, we have at present only the most meager data concerning the orderliness and richness of the child's syntactic input. Facts about the cross-linguistic similarities in the syntax/semantics correspondances are also extremely fragmentary, currently.

There are in addition numerous problems with the analyses performed that I have altogether skirted so far. For example, it is not an easy task to decide which structures co-occurring with verbs should actually be considered part of the frame

specifications, and which are merely adjuncts. To construct Table 2 (and in Lederer's ongoing work) we had to make some choices, but some of them may be wrong. And if we had these problems in assigning structural descriptions to the mother's utterances, isn't the learner similarly beset?[15] Another huge problem is the "idiomatic" verb uses that I mentioned in passing (footnote 10), e.g., John saw his victim out of the room, looked his enemies in the eye, etc. It may be significant that these monstrosities are just about totally absent from the maternal corpora we have examined, but absence in fact (rather than in principle) is a pretty weak reed on which to build so strong a position as the one I've tried to defend.

The largest problem of all is how learners acquire the semantic/syntactic linking rules in the first place. Bowerman's evidence, and all the findings I've just discussed, are understandable only (so far as I can see) by asserting that learners are in possession of such linking rules. But where did they come from? In the present discussion, I've subscribed to a version of Jackendoff's hypothesis that the linking rules are somehow cognitively transparent to the child. But since there is at least some cross-linguistic variance in such syntactic/semantic regularities (see Talmy, 1985), I admit that I'd be happier to find that they could be derived from some more primitive categories or functions. The problems here cry out for serious investigation.

In light of the various issues just mentioned, one must remain agnostic about the bootstrapping proposals, at present. But I hope I've persuaded you that the prospects they open for explanation of the verb-learning feat are enticing enough to make continued investigation seem worthwhile.

It remains to point out that, by their nature, both semantic and syntactic bootstrapping are perilous and errorful procedures and their explanatory power must be evaluated with this additional proviso in mind. Bowerman's children, drawing syntactic conclusions from meaningful overlap, are sometimes wrong. Errors are made insofar as the scenes are multiply interpretable; for instance, youngsters often interchange win

---

[15] There is some evidence in the literature of adult speech perception that adjunct and argument phrases may be intonationally distinguishable (see Gleitman and Wanner, 1982, for a review; and Carlson and Tannenhaus, 1988, for some experimental evidence). These distinctions, if real, can be expected to be exaggerated in maternal speech. Nevertheless, the issues here are quite complex and have not been thoroughly studied by any means. And they do bear in serious ways on the amount of work that syntax can be expected to do for the verb learner.

and beat, presumably because these occur in exactly the same circumstances. But syntactic bootstrapping is no more free of potential error. This is because the form-to-meaning mapping in the exposure language is complex and often inexact. For instance, exit, enter, reach and touch differ from most verbs describing directed motion through space in not requiring prepositional phrases to express the motion paths (compare come into the room but enter the room). One outcome of this inexact mapping of form onto meaning is errorful learning (e.g., the child may say "I touched on your arm") and its end point, language change (e.g., while exit the stage was the more common in Shakespeare's time, exit from the stage is now on the ascendancy). Short of changing the language, how do learners recover from such errors?

The position I have been urging is that children ferret out the forms and the meanings of the language just because they can play off these two imperfect and insufficient databases (the saliently interpreted events, and the syntactically interpreted utterances) against each other to derive the best fit between them. Neither syntactic nor semantic bootstrapping work all the time, nor taken together do they answer to all the questions about how children acquire their verb vocabulary. But I hope I've convinced you that each of these procedures works very well indeed when it does work, so the wise child should, and probably does, make use of both of them.

- - - 0 - - -

* This paper is the text of the keynote address delivered to the Stanford Child Languge Conference in April of 1989. The ideas contained in it were developed in collaboration with a number of colleagues and students, whose contributions are cited throughout the text. I am particularly indebted to two individuals who helped me throughout the preparation of this address. The first is my husband, Henry Gleitman, who -- as always -- quietly contributed a large share of the ideas and most of whatever organization and coherence this draft contains. Anne Lederer has also been a crucial aid in offering significant ideas and helping me get my head together on some of what's said here. I should add that, beyond their intellectual labors on my behalf, these colleagues were repeatedly willing to cut and paste, and even run and fetch, to help me meet deadlines. For both kinds of contribution, I am very grateful. I want also to express appreciation for a University of Pennsylvania Biomedical Research Grant (sponsored by the National Institute of Health under Grant # 2-S07-RR-07803-23) which underwrote the more recent experimental work that I report here.

REFERENCES

Acredolo, L. and Evans, D. (1980). Developmental changes in the effects of landmarks on infant spatial behavior Developmental Psychology 16: 312-318.

Baillargeon, R., Spelke, E. and Wasserman, S. (1985). Object permanence in 5-month old infants, Cognition 20 191-208.

Ball, W. and Vurpillot, E. (1976). Perception of movement in depth in infancy, L'Annee Psychologique 76:383- 91.

Berwick, R.C. (1981). Computational complexity and lexical functional grammar. In Proceedings of the 19th meeting of the Association for Computational Linguistics, Stanford, CA, 7 - 12.

Bowerman, M. (1976) Semantic factors in the acquisition of rules for word use and sentence construction. In D.M. Morehead & A.E. Morehead (Eds.) Normal and deficient child language. Baltimore: University Park Press.

Bowerman, M. (1982) Evaluating competing linguistic models with language acquisition data: Implications of developmental errors. Quaderni di Semantica. III, 5-66.

Bornstein, M.H. (1975) Qualities of color vision in infancy Journal of Experimental Child Psychology 19:401 - 419.

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

Bruner, J.S (1974/5) From communication to language: a psychological perspective. Cognition 3: 255-87.

Carey, S. (1982) Semantic development: The state of the art. In E. Wanner & L. Gleitman (Eds.) Language acquisition: The state of the art. New York: Cambridge University Press.

Carlson, G.N. & Tannenhaus, M.K. (1988) Thematic roles and language comprehension. In W. Wilkins (Ed.) Syntax and semantics 21: Thematic relations. New York: Academic Press.

Chomsky, N. (1959) Review of Skinner's Verbal behavior. Language 35:1 26 - 58.

Chomsky, N. (1975) Reflections on language New York: Random House.

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

Chomsky, N. (1982) Noam Chomsky on The Generative Enterprise: a discussion with R. Hybregts & H. van Riemsdijk. Holland: Dordrecht.

Crain, S. and Fodor, J.D. (in press), In E. Dromi (ed.), Language and cognition: A developmental perspective Norwood, N.J.: Ablex.

Fernald, A. (1984). The perceptual and affective salience of mothers' speech to infants. In L. Feagans, C. Garvey, & R. Golinkoff (Eds.). The origins and growth of communication. New Brunswick, N.J.: Ablex.

Field, J. (1976) Relation of young infant's reaching to stimulus distance and solidity. Child Development 50:698-704.

Fisher,C., Gleitman, H. and Gleitman, L. Relationships between verb meanings and their syntactic structures, Cognitive Psychology (in press)

Gibson, E.J. and E. Spelke (1983) The development of perception In J.H. Flavell & E. Markman (Eds), Cognitive development Vol III of Mussen (ed) Handbook of cognitive psychology New York: Wiley.

Gibson, E.J. and Walker, A.S (1984) Intermodal perception of substance. Child Development 55: 453-60.

Gleitman, L., Landau, B., Wanner, E., & Gleitman, H. (1987) Where learning begins: Initial representations for language learning. In E. Newmayer (Ed.) The Cambridge Linguistic Survey, Vol. II. New York: Cambridge University Press.

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

Gold, E.M. (1967). Language identification in the limit. Information and Control 10: 447-74.

Golinkoff, R., Harding, C., Carlson, V. & Sexton, E (1984) The child's perception of causal events: the distinction between animate and inanimate objects. IN L. P. Lipsitt & R. C. Rovee-Collier (Eds.) Advances in Infancy Research 3: 145-65.

Golinkoff, R., Hirsh-Pasek, P., Cauley, K. & Gordon, L. (1987) The eyes have it: lexical and syntactic comprehension in a new paradigm. Journal of Child Language.

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

Hirsh-Pasek, K., Goldinkoff, R., Fletcher, A., DeGaspe Beaubien, F. & Cauley, K. (1985) In the beginning: One word speakers comprehend word order. Paper presented at Boston Language Conference, October 1985.

Hirsh-Pasek, K., Kemler-Nelson, D.G., Jusczyk, P.W., Wright, K. & Druss, B. (1988) Clauses are perceptual units for young infants, Cognition

Hirsh-Pasek, K., Gleitman, H. , Gleitman, L.R., Golinkoff, R. & Naigles, L. (1988), Syntactic bootstrapping: Evidence from comprehension. Paper presented at Boston Language Conference, October 21, 1988.

Jackendoff, R. (1978) Grammar as evidence for conceptual structure. In M. Halle, J. Bresnan, and G. Miller (Eds.), Linguistic theory and psychological reality Cambridge, MA: The MIT Press.

Jackendoff, R. (1983) Semantics and cognition. Cambridge, MA: The MIT Press.

Joshi, A.K and Levy, L.S. (1982) Phrase structure trees bear more fruit than you would have thought, American Journal of Computational Linguistics 8:1, 1-12.

Katz, N., Baker, E. & MacNamara, J. (1974) What's in a name? A study of how children learn common and proper names. Child Development 45: 469-473.

Keil, F. C. (1979). Semantic and conceptual development. Cambridge: Harvard University Press.

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

Lasky, R.E. and Gogol, W.C. (19 8). The perception of relative motion by young infants. Perception 7, 617 - 623.

Lederer, A., Gleitman, L., and Gleitman, H. (1989) Input to a deductive, principle-driven, verb acquisition procedure. Speech delivered at the Boston Language Conference, October, 1989.

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

Marantz, A. (1982) On the acquisition of grammatical relations. Linguistische Berichte 80/82, 32 - 69.

Morgan, J., Meier, R. & Newport, E.L. (1988) Structural packaging in the input to language learning: contributions of intonational and morphological marking of phrases to the acquisition of language. Cognitive Psychology.

Morgan, J. & Newport, E.L. (1981) The role of constituent structure in the induction of an artificial language. Journal of Verbal Learning and Verbal Behavior 20: 67-85.

Naigles, L.G. (in press) Children use syntax to learn verb meanings, Journal of Child Language.

Naigles, L.G., Gleitman, L., and Gleitman, H. (in press) Children acquire word meaning components from syntactic evidence. In E. Dromi (Ed.), Language and cognition: a developmental perspective Norwood, N.J.: Ablex.

Newport, E.L. (1977) Motherese: The speech of mothers to young children. In N. Castellan, D. Pisoni & G. Potts (Eds.) Cognitive theory, Vol II, Hillsdale, N.J.: Erlbaum.

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

Pinker, S. (1987) Resolving a learnability paradox in the acquisition of the verb lexicon. Lexicon project working papers 17. Cambridge, MA: MIT Center for Cognitive Science.

Rescorla, L. (1980) Overextension in early language development Journal of Child Language 7.2 321-336.

Slobin, D.I. (1975) On the nature of talk to children. In E. H. Lenneberg & E. Lenneberg (Eds.), Foundations of language development, Vol I. New York: Academic Press.

Slobin, D.I. and Bever, T.G. (1982) Children use canonical sentence schemas: a crosslinguistic study of word order and inflections. Cognition 12 229-265.

Spelke, E.S. (1979) Perceiving bimodally specified events in infancy. Developmental Psychology 25:626-636.

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, N.J.: Erlbaum.

Starkey, P. Spelke, E.S. & Gelman, R. (1983) Detection of 1-1

correspondences by human infants. Science 210: 1033-5.

Streeter, L.A. (1978) Acoustic determinants of phrase boundary perception. Journal of the Acoustic Society of America 64: 1582-1592.

Talmy, L. (1975) Semantics and syntax of motion. In J. Kimball (Ed.), Syntax and semantics 4. New York: Academic Press.

Talmy, L. (1985) Lexicalization patterns: semantic structure in lexical forms. IN T. Shopen (Ed.) Language typology and language description. Cambridge: Cambridge University Press.

Wexler, K. and Culicover, P. (1980) Formal principles of language acquisition, Cambridge, MA: MIT Press.

Wexler, K. and Manzini, R. (1987) Parameters and learnability in binding theory. In Roeper and Williams (Eds.) Parameter Setting, Dordrecht: Reidel.

Zwicky, A. (1971) In a manner of speaking. Linguistic Inquiry, 11:2, 223-233.