AUTOMATIC ACQUISITION OF THE LEXICAL SEMANTICS OF VERBS FROM SENTENCE FRAMES*

Mort Webster and Mitch Marcus
Department of Computer and Information Science
University of Pennsylvania
200 S. 33rd Street
Philadelphia, PA 19104

ABSTRACT

This paper presents a computational model of verb acquisition which uses what we will call the principle of *structured overcommitment* to eliminate the need for negative evidence. The learner escapes from the need to be told that certain possibilities cannot occur (i.e., are "ungrammatical") by one simple expedient: It assumes that all properties it has observed are either obligatory or forbidden until it sees otherwise, at which point it decides that what it thought was either obligatory or forbidden is merely optional. This model is built upon a classification of verbs based upon a simple three-valued set of features which represents key aspects of a verb's syntactic structure, its predicate/argument structure, and the mapping between them.

1 INTRODUCTION

The problem of how language is learned is perhaps the most difficult puzzle in language understanding. It is necessary to understand learning in order to understand how people use and organize language. To build truly robust natural language systems, we must ultimately understand how to enable our systems to learn new forms themselves.

Consider the problem of learning new lexical items in context. To take a specific example, how is it that a child can learn the difference between the verbs *look* and *see* (inspired by Landau and Gleitman(1985))? They clearly have similar core meanings, namely "perceive by sight". One initially attractive and widely-held hypothesis is that

word meaning is learned directly by observation of the surrounding non-linguistic context. While this hypothesis ultimately only begs the question, it also runs into immediate substantive difficulties here, since there is usually looking going on at the same time as seeing and vice versa. But how can one learn that these verbs differ in that *look* is an active verb and *see* is stative? This difference, although difficult to observe in the environment, is clearly marked in the different syntactic frames the two verbs are found in. For example, *see*, being a stative perception verb, can take a sentence complement:

(1) John saw that Mary was reading.

while *look* cannot:

(2) * John looked that Mary was reading.

Also *look* can be used in an imperative,

(3) Look at the ball!

while it sounds a bit strange to command someone to *see*,

(4) ? See the ball!

(Examples like "look Jane, see Spot run!" notwithstanding.) This difference reflects the fact that one can command someone to direct their eyes (*look*) but not to mentally perceive what someone else perceives (*see*). As this example shows, there are clear semantic differences between verbs that are reflected in the syntax, but not obvious by observation alone. The fact that children are able to correctly learn the meanings of *look* and *see*, as well as hundreds of other verbs, with minimal exposure suggests that there is some correlation between syntax and semantics that facilitates the learning of word meaning.

Still, this and similar arguments ignore the fact that children do not have access to the negative

evidence crucial to establishing the active/stative distinction of the look/see pair. Children cannot know that sentences like (2) and (4) do not occur, and it is well established that children are not corrected for syntactic errors. Such evidence renders highly implausible models like that of Pinker(1987), which depend crucially on negative examples. How then can this semantic/syntactic correlation be exploited?

2 STRUCTURED OVERCOMMITMENT AND A LEARNING ALGORITHM

In this paper, we will present a computational model of verb acquisition which uses what we will call the principle of *structured overcommitment* to eliminate the need for such negative evidence. In essence, our learner learns by initially jumping to the strongest conclusions it can, simply assuming that everything within its descriptive system that it hasn't seen will never occur, and then later weakening its hypotheses when faced with contradictory evidence. Thus, the learner escapes from the need to be told that certain possibilities cannot occur (i.e. are "ungrammatical") by the simple expedient of assuming that all properties it has observed are either always obligatory or always forbidden. If and when the learner discovers that it was wrong about such a strong assumption, it reclassifies the property from either obligatory or forbidden to merely optional.

Note that this learning principal requires that no intermediate analysis is ever abandoned; analyses are only further refined by the weakening of universals (X ALWAYS has property P) to existentials (X SOMETIMES has property P). It is in this sense that the overcommitment is "structured." For such a learning strategy to work, it must be the case that the set of features which underlies the learning process are surface observable; the learner must be able to determine of a particular instance of (in this case) a verb structure whether some property is true or false of it. This would seem to imply, as far as we can tell, a commitment to the notion of em learning as selection widely presupposed in the linguistic study of generative grammar (as surveyed, for example, in Berwick(1985). Thus, we propose that the problem of learning the category of a verb does not require that a natural language understanding system synthesize em de novo a new structure to represent its semantic class, but rather that it determine to which of

a predefined, presumably innate set of verb categories a given verb belongs. In what follows below, we argue that a relevant classification of verb categories can be represented by simple conjunctions of a finite number of predefined quasi-independent features with no need for disjunction or complex boolean combinations of features.

Given such a feature set, the Principal of Structured Overcommitment defines a partial ordering (or, if one prefers, a tangled hierarchy) of verbs as follows: At the highest level of the hierarchy is a set of verb classes where all the primary four features, where defined, are either obligatory or forbidden. Under each of these "primary" categories there are those categories which differ from it only in that some category which is obligatory or forbidden in the higher class is optional in the lower class. Note that both obligatory and forbidden categories at one level lead to the same optional category at the next level down.

The learning system, upon encountering a verb for the first time, will necessarily classify that verb into one of the ten top-level categories. This is because the learner assumes, for example, that if a verb is used with an object upon first encounter, that it always has an object; if it has no object, that it never has an object, etc. The learner will leave each verb classification unchanged upon encountering new verb instances until a usage occurs that falsifies at least one of the current feature values. When encountering such a usage i.e. a verb frame in which a property that is marked obligatory is missing, or a property that is marked forbidden is present (there are no other possibilities) - then the learner reclassifies the verb by moving down the hierarchy at least one level replacing the OBLIGATORY or FORBIDDEN value of that feature with OPTIONAL.

Note that, for each verb, the learner's classification moves monotonically lower on this hierarchy, until it eventually remains unchanged because the learner has arrived at the correct value. (Thus this learner embodies a kind of em learning in the limit.

3 THE FEATURE SET AND THE VERB HIERARCHY

As discussed above, our learner describes each verb by means of a vector of features. Some of these features describe syntactic properties of the verb (e.g. "Takes an Object"), others describe aspects of the theta-structure (the predicate/argument structure) of the verb (e.g. "Takes

an Agent","Takes a Theme"), while others describe some key properties of the mapping between theta-structure and syntactic structure (e.g. "Theme Appears As Surface Object"). Most of these features are three-valued; they describe properties that are either always true (e.g. that "devour" always Takes An Object), always false (e.g. that "fall" never Takes An Object) or properties that are optionally true (e.g. that "eat" optionally Takes An Object). Always true values will be indicated as "+" below, always false values as "−" and optional values as "0".

All verbs are specified for the first three features mentioned above: "Takes an Object" (OBJ), "Takes an Agent" (AGT), and "Takes a Theme" (THEME). All verbs that allow OBJ and THEME are specified for "Theme Appears As Object" (TAO), otherwise TAO is undefined. At the highest level of the hierarchy is a set of verb classes where all these primary features, where defined, are either obligatory or forbidden. Thus there are at most 10 primary verb types; of the eight for the first three features, only two (+−+, and +++) split for TAO.

The full set of features we assume include the primary set of features (OBJ, AGT, THEME, and TAO), as described above, and a secondary set of features which play a secondary role in the learning algorithm, as will be discussed below. These secondary features are either thematic properties, or correlations between thematic and syntactic roles. The thematic properties are: LOC - takes a locative; INST - takes an instrument; and DAT - takes a dative. The first thematic-syntactic mapping feature "Instrument as Subject" is false if no instrument can appear in subject position (or, true if the subject is always an instrument, although this is never the case.) The second such feature "Theme as Chomeur" (TAC) is the only non-trinary-valued feature in our learner; it specifies what preposition marks the theme when it is not realized as subject or object. This feature, if not −, either takes a lexical item (a preposition, actually, as its value, or else the null string. We treat verbs with double objects (e.g. "John gave Mary the ball.") as having a Dative as object, and the theme as either marked by a null preposition or, somewhat alternatively, as a bare NP chomeur. (The facts we deal with here don't decide between these two analyses.)

Note that this analysis does not make explicit what can appear as object; it is a claim of the analysis that if the verb is OBJ:+ or OBJ:0 and is TAO:− or TAO:0, then whatever other thematic roles may occur can be realized as the object. This may well be too strong, but we are still seeking a counterexample.

Figure 1 shows our classification of some verb classes of English, given this feature set. (This classification owes much to Levin(1985), as well as to Grimshaw(1983) and Jackendoff(1983).) This is only the beginning of such a classification, clearly; for example, we have concentrated our efforts solely on verbs that take simple NPs as complements. Our intention is merely to provide a rich enough set of verb classes to show that our classification scheme has merit, and that the learning algorithm works. We believe that this set of features is rich enough to describe not only the verb classes covered here but other similar classes. It is also our hope that an analysis of verbs with richer complement structures will extend the set of features without changing the analysis of the classes currently handled.

It is interesting to note that although the partial ordering of verb classes is defined in terms of features defined over syntactic and theta structures, that there appears to be at least a very strong semantic reflex to the network. Due to lack of space, we label verb classes in Figure 1 only with exemplars; here we give a list of either typical verbs in the class, and/or a brief description of the class, in semantic terms:

- Spray, load, inscribe, sow: Verbs of physical contact that show the completive/noncompletive[1] alternation. If completive, like "fill".

- Clear, empty: Similar to spray/load, but if completive, like "empty".

- Wipe: Like clear, but no completive pattern.

- Throw: The following four verb classes all involve an object and a trajectory. "Throw" verbs don't require a terminus of the trajectory.

- Present: Like "throw", as far as we can tell.

- Give: Requires a terminus.

---

[1] This is the difference between:

   I loaded the hay on the truck.

and

   I loaded the truck with hay.

In the second case, but not the first, there is a implication that the truck is completely full.

	OBJ	AGT	THEME	TAO	LOC	INST	DAT	TAC	INST AS SUBJ
SPRAY, LOAD	+	+	0	0	0	0	−	with	0
CLEAR, EMPTY	+	+	0	0	+	0	−	of	−
WIPE	+	+	0	0	+	0	−	−	−
PRESENT	+	+	+	0	−	0	0	with	0
SEARCH	0	+	0	−	0	0	−	for	−
MELT, BREAK	0	0	+	0	−	0	−	−	0
PIERCE, DESTROY	+	0	+	+	−	0	−	−	0
POKE, TOUCH	+	+	+	0	0	0	−	−	0
PUT	+	+	+	+	+	0	−	−	0
DEVOUR DYNAMITE	+	+	+	+	−	−	−	−	−
HUG	+	+	+	+	0	0	−	−	−
EAT	0	+	0	0	−	0	−	−	−
SWIM, FLY	−	+	−	−	0	−	−	−	−
BREATHE	0	+	0	0	0	−	−	−	−
FILL	+	+	0	−	+	0	−	with	0
THROW	+	+	+	0	−	−	0	<null>	−
PUSH	0	+	+	0	0	0	−	at against	0
GIVE	+	+	+	0	−	−	+	<null>	−
STAND	+	0	+	+	0	0	−	−	−
DIE, FLOWER	−	−	+	−					
RAIN	−	−	0	−					

Figure 1: Some verb feature descriptions.
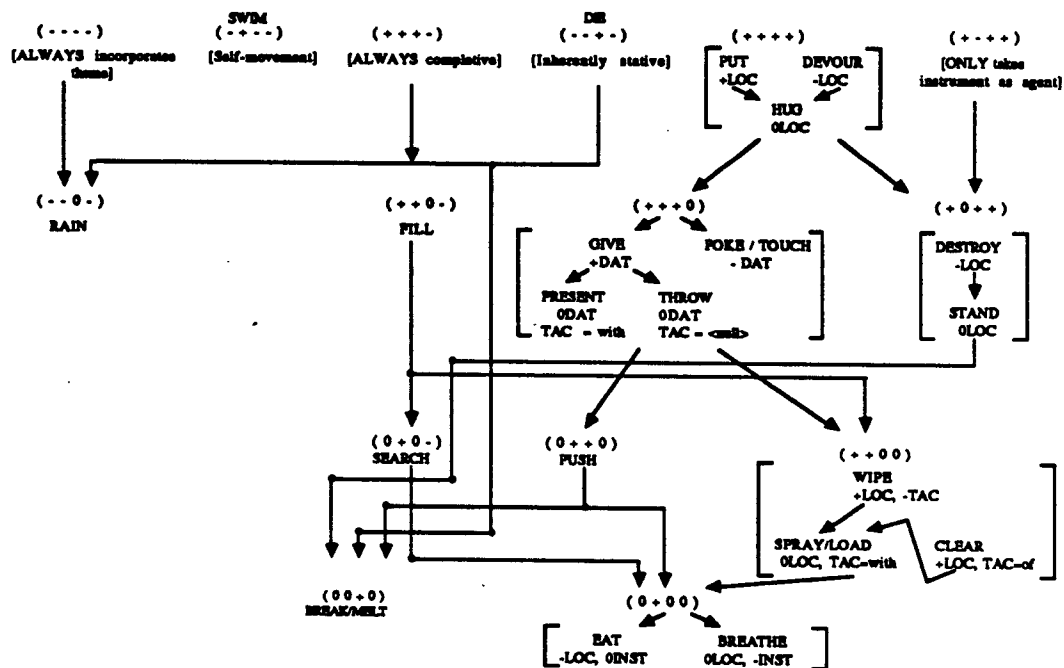

Figure 2: The verb hierarchy.

- Poke, jab, stick, touch: Some object follows a trajectory, resulting in surface contact.

- Hug: Surface contact, no trajectory.

- Fill: Inherently completive verbs.

- Search: Verbs that show a completive/non-completive alternation that doesn't involve physical contact.

- Die, flower: Change of state. Inherently non-agentive.

- Break: Change of state, undergoing causitive alternation.

- Destroy: Verbs of destruction.

- Pierce: Verbs of destruction involving a trajectory.

- Devour, dynamite: Verbs of destruction with incorporated instruments

- Put: Simple change of location.

- Eat: Verbs of ingesting allowing instruments

- Breathe: Verbs of ingesting that incorporate instrument

- Fall, swim: Verbs of movement with incorporated theme and incorporated manner.

- Push: Exerting force; maybe something moves, maybe not.

- Stand: Like "break", but at a location.

- Rain: Verbs which have no agent, and incorporate their patient.

The set of verb classes that we have investigated interacts with our learning algorithm to define the partial order of verb classes illustrated schematically in Figure 2.

For simplicity, this diagram is organized by the values of the four principle features of our system. Each subsystem shown in brackets shares the same principle features; the individual verbs within each subsystem differ in secondary features as shown. If one of the primary features is made optional, the learning algorithm will map all verbs in each subsystem into the same subordinate subsystem as shown; of course, secondary feature values are maintained as well. In some cases, a sub-hierarchy within a subsystem shows the learning of a secondary feature.

We should note that several of the primary verb classes in Figure 2 are unlabelled because they correspond to no English verbs: The class "————" would be the class of rain if it didn't allow forms like "hail stones rained from the sky", while the class "+−++" would be the class of verbs like "destroy" if they only took instruments as subjects. Such classes may be artifacts of our analysis, or they may be somewhat unlikely classes that are filled in languages other than English.

Note that sub—patterns in the primary feature subvector seem to signal semantic properties in a straightforward way. So, for example, it appears that verbs have the pattern {OBJ:+, THEME:+, TAO:−} only if they are inherently completive; consider "search" and "fill". Similarly, the rare verbs that have the pattern {OBJ:−, THEME:−}, i.e those that are truly intransitive, appear to incorporate their theme into their meaning; a typical case here is "swim". Verbs that are {OBJ:−, AGT:−} (e.g. "die") are inherently stative; they allow no agency. Those verbs that are {AGT:+} incorporate the instrument of the operation into their meaning. We will have to say about this below.

4 THE LEARNING ALGORITHM AT WORK

Let us now see how the learning algorithm works for a few verbs.

Our model presupposes that the learner receives as input a parse of the sentence from which to derive the subject and object grammatical relations, and a representation of what NPs serve as agent, patient, instrument and location. This may be seen as begging the question of verb acquisition, because, it may be asked, how could an intelligent learner know what entities function as agent, patient, etc. without understanding the meaning of the verb? Our model in fact presupposes that a learner can distinguish between such general categories as animate, inanimate, instrument, and locative from direct observation of the environment, without explicit support from verb meaning; i.e. that it will be clear from observation em who is acting on em what em where. This assumption is not unreasonable; there is strong experimental evidence that children do in fact perceive even something as subtle as the difference between animate and inanimate motion well before the two word stage (see Golinkoff et al, 1984). This notion that agent, patient and the like can be derived from direct observation (perhaps focussed by what NPs

appear in the sentence) is a weak form of what is sometimes called the em semantic bootstrapping hypothesis (Pinker(1984)). The theory that we present here is actually a combination of this weak form of semantic bootstrapping with what is called em syntactic bootstrapping, the notion that syntactic frames alone offer enough information to classify verbs (see Naigles, Gleitman, and Gleitman (in press) and Fisher, Gleitman and Gleitman(1988).)

With this preliminary out of the way, let's turn to a simple example. Suppose the learner encounters the verb "break", never seen before, in the context

(6) The window broke.

The learner sees that the referent of "the window" is inanimate, and thus is the theme. Given this and the syntactic frame of (6), the learner can see that em break (a) does not take an object, in this case, (b) does not take an agent, and (c) takes a patient. By Structured Overcommitment, the learner therefore assumes that em break em never takes an object, em never takes a subject, and em always takes a patient. Thus, it classifies em break as {OBJ:−, AGT:−, THEME:+, TAO:−} (if TAO is undefined, it is assigned "−"). It also assumes that em break is {DAT:−, LOC:−, INST:−, ...} for similar reasons. This is the class of DIE, one of the toplevel verb classes.

Next, suppose it sees

(7) John broke the window.

and sees from observation that the referent of "John" is an agent, the referent of "the window" a patient, and from syntax that "John" is subject, and "the window" object. That em break takes an object conflicts with the current view that em break NEVER takes an object, and therefore this strong assumption is weakened to say that em break SOMETIMES takes an object. Similarly, the learner must fall back to the position that em break SOMETIMES can have the theme serve as object, and can SOMETIMES have an agent. This takes {OBJ:−, AGT:−, THEME:+, TAO:−} to {OBJ:0, AGT:0, THEME:+, TAO:0}, which is the class of both em break and em stand. However, since it has never seen a locative for em break, it assumes that em break falls into exactly the category we have labelled as "break".[2]

[2] And how would it distinguish between

The vase stood on the table.

and

There are, of course, many other possible orders in which the learner might encounter the verb em break. Suppose the learner first encounters the pattern

(8) John broke the window.

before any other occurrences of this verb. Given only (8), it will assume that em break always takes an object, always takes an agent, always has a patient, and always has the patient serving as object. The learner will also assume that em break never takes a location, a dative, etc. This will give it the initial description of {OBJ:+, AGT:+, THEME:+, TAO:+, ..., LOC:−}, which causes the learner to classify em break as falling into the toplevel verb class of DEVOUR, verbs of destruction with the instrument incorporated into the verb meaning.

Next, suppose the learner sees

(9) The hammer broke the window.

where the learner observes that "hammer" is an inanimate object, and therefore must serve as instrument, not agent. This means that the earlier assumption that agent is necessary was an overcommitment (as was the unmentioned assumption that an instrument was forbidden). The learner therefore weakens the description of em break to {OBJ:+, AGT:0, THEME:+, TAO:+, ..., LOC:−, INST:0}, which moves em break into the verb class of DESTROY, destruction without incorporated instrument.

Finally (as it turns out), suppose the learner sees

(10) The window broke.

Now it discovers that the object is not obligatory, and also that the theme can appear as subject, not object, which means that TAO is optional, not obligatory. This now takes em break to {OBJ:0, AGT:0, THEME:+, TAO:0, ...}, which is the verb class of break.

We interposed (9) between (8) and (10) in this sequence just to exercise the learner. If (10) followed (8) directly, the learner would have taken em break to verb class BREAK all the more quickly. Although we will not explicitly go through the exercise here, it is important to our claims that any permutation of the potential sentence frames of em break will take the learner to BREAK, although some combinations require verb classes not shown

The base broke on the table?

This is a problem we discuss at the end of this paper.

on our chart for the sake of simplicity (e.g. the class {OBJ:0, AGT:−, THEME:+, TAO:0} if it hasn't yet seen an agent as subject.).

We were somewhat surprised to note that the trajectory of em break takes the learner through a sequence of states whose semantics are useful approximations of the meaning of this verb. In the first case above, the learner goes through the class of "change of state without agency", into the class of BREAK, i.e. "change of state involving no location". In the second case, the trajectory takes the learner through "destroy with an incorporated instrument", and then DESTROY into BREAK. In both of these cases, it happens that the trajectory of em break through our hierarchy causes it to have a meaning consistent with its final meaning at each point of the way. While this will not always be true, it seems that it is quite often the case. We find this property of our verb classification very encouraging, particularly given its genesis in our simple learning principle.

We now consider a similar example for a different verb, the verb em load, in somewhat terser form. And again, we have chosen a somewhat indirect route to the final derived verb class to demonstrate complex trajectories through the space of verb classes. Assume the learner first encounters

(11) John loads the hay onto the truck.

From (11), the learner builds the representation {OBJ:+, AGT:+, THEME:+, TAO:+, ..., LOC:+, ..., DAT:−}, which lands the learner into the class of PUT, i.e. "simple change of location". We assume that the learner can derive that "the truck" is a locative both from the prepositional marking, and from direct observation.

Next the learner encounters

(12) John loads the hay.

From this, the learner discovers that the location is not obligatory, but merely optional, shifting it to {OBJ:+, AGT:+, THEME:+, TAO:+, ..., LOC:0 ..., DAT:−}, the verb class of HUG, with the general meaning of "surface contact with no trajectory."

The next sentence encountered is

(13) John loads the truck with hay.

This sentence tells the learner that the theme need only optionally serve as object, that it can be · shifted to a non-argument position marked with the preposition em with. This gives em load the description of {OBJ:+, AGT:+, THEME:+,

TAO:0, TAC:with, ..., LOC:0 ..., DAT:−}. This new description takes em load now into the verb class of POKE/TOUCH, surface contact by an object that has followed some trajectory. (We have explicitly indicated in our description here that {DAT:−} was part of the verb description, rather than leaving this fact implicit, because we knew, of course, that this feature would be needed to distinguish between the verb classes of GIVE and POKE/TOUCH. We should stress that this and many other features are encoded as "−" until encountered by the learner; we have simply suppressed explicitly representing such features in our account here unless needed.)

Finally, the learner encounters the sentence

(14) John loads the truck.

which makes it only optional that the theme must occur, shifting the verb representation to {OBJ:+, AGT:+, THEME:0, TAO:0, TAC:with, ..., LOC:0 ..., DAT:−}. The principle four features of this description put the verb into the general area of WIPE, CLEAR and SPRAY/LOAD, but the optional locative, and the fact that the theme can be marked with em with select for the class of SPRAY/LOAD, verbs of physical contact that show the completive/noncompletive alternation."

Note that in this case again, the semantics of the verb classes along the learning trajectory are reasonable successive approximations to the meaning of the verb.

5 FURTHER RESEARCH AND SOME PROBLEMS

One difficulty with this approach which we have not yet confronted is that real data is somewhat noisy. For example, although it is often claimed that Motherese is extremely clean, one researcher has observed that the verb "put", which requires both a location and an object to be fully grammatical, has been observed in Motherese (although extremely infrequently) without a location. We strongly suspect, of course, that the assumption that one instance suffices to change the learner's model is too strong. It would be relatively easy to extend the model we give here with a couple of bits to count the number of counterexamples seen for each obligatory or forbidden feature, with two or three examples needed within some limited time period to shift the feature to optional.

Can the model we describe here be taken as a psychological model? At first glance, clearly not,

183

because this model appears to be deeply conservative, and as Pinker(1987) demonstrates, children freely use verbs in patterns that they have not seen. In our terms, they use verbs as if they had moved them down the hierarchy without evidence. The facts as currently understood can be accounted for by our model given one simple assumption: While children summarize their exposure to verb usages as discussed above, they will use those verbs in highly productive alternations (as if they were in lower categories) for some period after exposure to the verb. The claim is that their em usage might be non-conservative, even if their representations of verb class are. By this model, the child would restrict the usage of a given verb to the represented usages only after some period of time. The mechanisms for deriving criteria for productive usage of verb patterns described by Pinker(1987) could also be added to our model without difficulty. In essence, one would then have a non-conservative learner with a conservative core.

REFERENCES

[1] Berwick, R. (1985) *The Acquisition of Syntactic Knowledge*. Cambridge, MA: MIT Press.

[2] Fisher, C.; Gleitman, H.; and Gleitman, L. (1988) Relations between verb syntax and verb semantics: On the semantic content of subcategorization frames. Submitted for publication.

[3] Golinkoff, R.M.; Harding, C.G.; Carson, V.; and Sexton, M.E. (1984) The infant's perception of causal events: the distinction between animate and inanimate object. In L.P. Lipsitt and C. Rovee-Collier (Eds.) *Advances in Infancy Research 3*: 145-65.

[4] Grimshaw, J. (1983) Subcategorization and grammatical relations. In A. Zaenen (Ed.), *Subjects and other subjects*. Evanston: Indiana University Linguistics Club.

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

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

[7] Levin, B. (1985) Lexical semantics in review: An introduction. In B. Levin (Ed.), Lexical semantics in review. *Lexicon Project Working Papers, 1*. Cambridge, MA: MIT Center for Cognitive Science.

[8] Naigles, L.; Gleitman, H.; and Gleitman, L.R. (in press) Children acquire word meaning components from syntactic evidence. In E. Dromi (Ed.) *Linguistic and conceptual development*. Ablex.

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

[10] 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.