On the Acquisition of Lexical Entries: The Perceptual Origin of Thematic Relations

James Pustejovsky

Department of Computer Science
Brandeis University
Waltham, MA 02254
617-736-2709
jamesp@brandeis.csnet-relay

Abstract

This paper describes a computational model of concept acquisition for natural language. We develop a theory of lexical semantics, the *Extended Aspect Calculus*, which together with a "markedness theory" for thematic relations, constrains what a possible word meaning can be. This is based on the supposition that predicates from the perceptual domain are the primitives for more abstract relations. We then describe an implementation of this model, TULLY, which mirrors the stages of lexical acquisition for children.

1. Introduction

In this paper we describe a computational model of concept acquisition for natural language making use of positive-only data, modelled on a theory of lexical semantics. This theory, the *Extended Aspect Calculus* acts together with a markedness theory for thematic roles to constrain what a possible word type is, just as a grammar defines what a well-formed tree structure is in syntax. We argue that linguistic specific knowledge and learning principles are needed for concept acquisition from positive evidence alone. Furthermore, this model posits a close interaction between the predicates of visual perception and the early semantic interpretation of thematic roles as used in linguistic expressions. In fact, we claim that these relations act as constraints to the development of predicate hierachies in language acquisition. Finally, we describe TULLY, an implementation of this model in *ZETALISP* and discuss its design in the context of machine learning research.

There has been little work on the acquisition of thematic relation and case roles, due to the absence of any consensus on their formal properties. In this research we begin to address what a theory of thematic relations might look like, using learnability theory as a metric for evaluating the model. We claim that there is an important relationship between visual or imagistic perception and the development of thematic relations in linguistic usage for a child. This has been argued recently by Jackendoff (1983, 1985) and was an assumption in the pioneering work of Miller and Johnson-Laird (1976). Here we argue that the conceptual abstraction of thematic information does not develop arbitrarily but along a given, predictable path; namely, a developmental path that starts with tangible perceptual predicates (e.g. spatial, causative) to later form the more abstract mental and cognitive predicates. In this view thematic relations are actually sets of thematic properties, related by a partial ordering. This effectively establishes a markedness theory for thematic roles that a learning system must adhere to in the acquisition of lexical entries for a language.

We will discuss two computational methods for concept development in natural language:

(1) *Feature Relaxation* of particular features of the arguments to a verb. This is performed by a constraint propagation method.

(2) *Thematic Decoupling* of semantically incorporated information from the verb.

When these two learning techniques are combined with the model of lexical semantics adopted here, the stages of development for verb acquisition are similar to those acknowledged for child language acquisition.

2. Learnability Theory and Concept Development

Work in machine learning has shown the usefulness to an inductive concept-learning system of inducing "bias" in the learning process (cf. [Mitchell 1977, 1978], [Michalski 1983]). An even more promising development is the move to base the bias on domain-intensive models, as seen in [Mitchell et al. 1985], [Utgoff 1985], and [Winston et al. 1983]. This is an important direction for those concerned with natural language acquisition, as it converges with a long-held belief of many psychologists and linguists that domain-specific information is necessary for learning (cf. [Slobin 1982], [Pinker 1984], [Bowerman 1974], [Chomsky 1980]). Indeed, Berwick (1984) moves in exactly this direction. Berwick describes a model for the acquisition of syntactic knowledge based on a restricted $\bar{X}$-syntactic parser, a modification of the Marcus parser ([Marcus 1980]). The domain knowledge specified to the system in this case is a parametric parser and learning system that adapts to a particular linguistic environment, given only positive data. This is just the sort of biasing necessary to account for data on syntactic acquisition.

One area of language acquisition that has not been sufficiently addressed within computational models is the acquisition of conceptual structure. For language acquisition, the problem can be stated as follows: How does the child identify a particular thematic role with a specific grammatical function in the sentence? This is the problem of mapping the semantic functions of a proposition into specified syntactic positions in a sentence.

Pinker (1984) makes an interesting suggestion (due originally to D. Lebeaux) in answer to this question. He proposes that one of the strategies available to the language learner involves a sort of "template matching" of argument to syntactic position. There are *canonical configurations* that are the default mappings and *non-canonical* mappings for the exceptions. For example, the template consists of two rows, one of thematic roles, and the other of syntactic positions. A canonical mapping exists if no lines joining the two rows cross. Figure 1 shows a canonical mapping representing the sentence in (1), while Figure 2 illustrates a noncanonical mapping representing sentence (2).

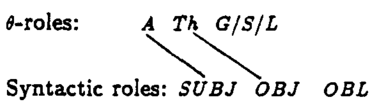


Figure 1

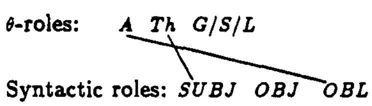


Figure 2

(1) Mary hit Bill.

(2) Bill was hit by Mary.

With this principle we can represent the productivity of verb forms that are used but not heard by the child. We will adopt a modified version of the canonical mapping strategy for our system, and embed it within a theory of how perceptual primitives help derive linguistic concepts.

As mentioned, one of the motivations for adopting the canonical mapping principle is the power it gives a learning system in the face of positive-only data. In terms of learnability theory, Berwick (1985) (following [Angluin 1978]) notes that to ensure successful acquisition of the language after a finite number of positive examples, something like the Subset Principle is necessary. We can compare this principle to a Version Space model of inductive learning([Mitchell 1977, 1978]), with no negative instances. Generalization proceeds in a conservative fashion, taking only the narrowest concept that covers the data.

How does this principle relate to lexical semantics and the way thematic relations are mapped to syntactic positions? We claim that the connection is very direct. Concept learning begins with spatial, temporal, and causal predicates being the most salient. This follows from our supposition that these are innate structures, or are learned very early. Following Miller and Johnson-Laird (1976), [Miller 1985], and most psychologists, we assume the prelinguistic child is already able to discern spatial orientations, causation, and temporal dependencies. We take this as a point of departure for our theory of markedness, which is developed in the next section.

3.0 Theoretical Assumptions

3.1 The Extended Aspect Calculus

In this section we outline the semantic framework which defines our domain for lexical acquisition. In the current linguistic literature on case roles or thematic relations, there is little discussion on what logical connection exists between one θ-role and another. Besides being the workhorse for motivating several principles of syntax (cf. [Chomsky 1981], [Williams 1980]) the most that is claimed is that Universal Grammar specifies a repertoire of thematic relations (or case roles), *Agent*, *Theme*, *Patient*, *Goal*, *Source*, *Instrument*, and that every NP must carry one and only one role. It should be remembered, however, that thematic relations were originally conceived in terms of the argument positions of semantic predicates such as CAUSE and DO. [1] That is a verb didn't simply have a list of labelled arguments [2] such as *Agent* and *Patient*, but had an interpretation in terms of more primitive predicates where the notions *Agent* and *Patient* were defined. The causer of an event (following Jackendoff (1976)) is defined as an *Agent*, for example, $CAUSE(x, e) \rightarrow Agent(x)$.

Similarly, the first argument position of the predicate GO is interpreted as Theme, as in $GO(x, y, z)$. The second argument here is the *SOURCE* and the third is called the *GOAL*.

The model we have in mind acts to constrain the space of possible word meanings. In this sense it is similar to Dowty's aspect calculus but goes beyond it in embedding his model within a markedness theory for thematic types. Our model is a first-order logic that employs symbols acting as special operators over the standard logical vocabulary. These are taken from three distinct semantic fields. They are: **causal**, **spatial**, and **aspectual**.

The predicates associated with the causal field are *Causer* (C_1), *Causee* (C_2), and *Instrument* (I). The spatial field has only one predicate, *Locative*, which is predicated of an object we term the *Theme*. Finally, the aspectual

[1] Cf.Jackendoff (1972, 1976) for a detailed elaboration of this theory.

[2] This is now roughly the common assumption in GB, GPSG, and LFG.

field has three predicates, representing the three temporal intervals t_1, beginning, t_2, middle, and t_3, end. From the interaction of these predicates all thematic types can be derived. We call the lexical specification for this aspectual and thematic information the *Thematic Mapping Index*.

As an example of how these components work together to define a thematic type, consider first the distinction between a state, an activity (or process), and an accomplishment. A state can be thought of as reference to an unbounded interval, which we will simply call t_2; that is, the state spans this interval.[3] An activity or process can be thought of as referring to a designated initial point and the ensuing process; in other words, the situation spans the two intervals t_1 and t_2. Finally, an event can be viewed as referring to both an activity and a designated terminating interval; that is, the event spans all three intervals, t_1, t_2, and t_3.

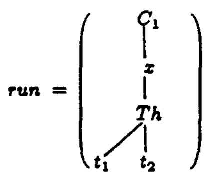
Now consider how these bindings interact with the other semantic fields for the verb *run* in sentence (8) and *give* in sentence (9).

(8) John ran yesterday.

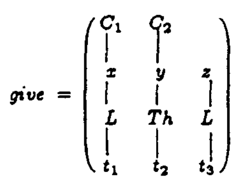
(9) John gave the book to Mary.

We associate with the verb *run* an argument structure of simply $run(x)$. For *give* we associate the argument structure $give(x, y, z)$. The Thematic Mapping Index for each is given below in (10) and (11).

(10)



(11)



The sentence in (8) represents a process with no logical culmination, and the one argument is linked to the named case role, *Theme*. The entire process is associated with both the initial interval t_1 and the middle interval t_2. The argument x is linked to C_1 as well, indicating that it is an *Actor* as well as a moving object (i.e. *Theme*). This represents one TMI for an activity verb.

The structure in (9) specifies that the meaning of *give* carries with it the supposition that there is a logical culmination to the process of giving. This is captured by reference to the final subinterval, t_3. The linking between x and the L associated with t_1 is interpreted as *Source*, while the other linked arguments, y and z are *Theme* (the book) and *Goal*, respectively. Furthermore, x is specified as a *Causer* and the object which is marked *Theme* is also an affected object (i.e. *Patient*). This will be one of the TMIs for an accomplishment.

In these examples the three subsystems are shown as rows, and the configuration given is lexically specified. [4]

[3] This is a simplication of our model, but for our purposes the difference is moot. A state is actually interpreted as a primitive homogeneous event-sequence, with downward closure. Cf. [Pustejovsky, 1987],

[4] [Jackendoff 1985] develops a similar idea, but *vide infra* for discussion.

3.2 A Markedness Theory for Thematic Roles

As mentioned above, the theory we are outlining here is grounded on the supposition that all relations in the language are suffiently described in terms of causal, spatial and aspectual predicates. A thematic role in this view is seen as a set of primitive properties relating to the predicates mentioned above. The relationship between these thematic roles is a partial ordering over the sets of properties defining them. It is this partial ordering that allows us to define a *markedness theory* for thematic roles. Why is this important?

If thematic roles are assigned randomly to a verb, then one would expect that there exist verbs that have only *Patient* or *Instrument*, or two *Agents* or *Themes*, for example. Yet this is not what we find. What appears to be the case is that thematic roles are not assigned to a verb independently of one another, but rather that some thematic roles are fixed only after other roles have been established. For example, a verb will not be assigned a *GOAL* if there is not a *THEME* assigned first. Similarly, a *LOCATIVE* is dependent on there being a *THEME* present. This dependency can be viewed as an acquisition strategy for learning the thematic relations of a verb.

Now let us outline the theory. We begin by establishing the most unmarked relation that an argument can bear to its predicate. Let us call this role $Theme_U$. The only semantic information this carries is that of an existential quantifier. It is the only named role outside of the three interpretive systems defined above. Normally, we think of *Theme* as an object in motion. This is only half correct, however, since statives carry a *Theme* readings as well. It is in fact the feature $[\pm motion]$ that distinguishes the role of Mary in (1) and (2) below.

(1) Stative: $[-motion]$ Mary sleeps.

(2) Active: $[+motion]$ Mary fell.

This gives us our first markedness convention:

(3) $Theme_U \rightarrow Theme_A / [+motion]$

(3) $Theme_U \rightarrow Theme_S / [-motion]$

where $Theme_A$ is an "activity" Theme, and $Theme_S$ is a stative.

Within the spatial subsystem, there is one variable type, *Location*, and a finite set of them $L_1, L_2 \ldots L_n$. The most unmarked location is that carrying no specific aspectual binding. That is, the named variables are L_B and L_E and are commonly referred to as *Source* and *Goal*. Thus, L_U is the unmarked role. The limitations on named locative variables is perhaps constrained only by the aspectual system of the language (rich aspectual distinction, then more named locative variables). The markedness conventions here are:

(4) $L_U \rightarrow S/B$

(5) $L_U \rightarrow G/E$

Within the causal subsystem there are three predicates, C_1, C_2, and I. We call C_2, (the traditional *Patient* role) is less marked than C_1, but is more marked than I.

These conventions give us the core of the primitive semantic relations. To be able to perform predicate generalization over each relation, however, we define a set of features that applies to each argument within the semantic subsystems. These are the abstraction operators that allow a perceptual-based semantics to generalize to non-perceptual relations. These features also have marked and unmarked values, as we will show below. There are four features that contribute to the generalization process in concept acquisition:

(a) [±*abstract*] (b) [±*direct*]

(c) [±*complete*] (d) [±*animate*]

The first feature, *abstract*, distinguishes tangible objects from intangible ones. *Direct* will allow a gradience in the notion of causation and motion. The third feature, *complete*, picks out the extension of an argument as either an entire object or only part of it. *Animacy* has the standard semantics of labeling an object as alive or not.

Let us illustrate how these operators abstract over primitive thematic roles. By changing the value of a feature, we can alter the description, and hence, the set of objects in its extension. Assume, for example, that the predicate C_1 has as its unmarked value, [+*Direct*].

(6) $C_1[U Direct] \rightarrow [+Direct]$

By changing the value of this feature we allow C_1, the direct agent of an event, to refer to an indirect causer.

(7) $Agent[+Direct] <_G Agent[-Direct]$

Similarly, we can change the value of the default setting for the feature [+*Complete*] to refer to a subcausation (or causation by part).

(8) $Agent[+Complete] <_G Agent[-Complete]$

These changes define a new concept, "effector", which is a superset of the previous concepts given in the system. The same can be done with C_2 to arrive at the concept of an "effected object." We see the difference in interpretation in the sentences below.

a. John intentionally broke the chair. (Agent-direct)

b. John accidentally broke that chair when he sat down. (Agent-indirect)

c. John broke the chair when he fell. (Effector)

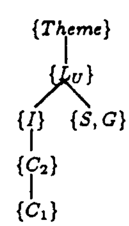
Given the manner in which the features of primitive thematic roles are able to change their values, we are defining a predictable generalization path that relations incorporating these roles will take. In other words, two concepts may be related thematically, but may have very different extensional properties. For example, *give* and *take* are clearly definable perceptual transfer relations. But given the abstractions available from our markedness theory, they are thematically related to something as distant as "experiencer verbs", e.g. *please*, as in "The book pleased John." This relation is a transfer verb with an *incorporated Theme*; namely, the "pleasure." [5]

If we apply these features in the spatial subsystem, we can arrive at generalized notions of location, as well as abstracted interpretations for *Theme*, *Goal* and *Source*. For example, given the thematic role $Th-A$ with the feature [−*Abstract*] in the default setting, we can generalize to allow for abstract relations such as *like*, where the object is not affected, but is an abstract Theme. Similarly, the *Theme* in a sentence such as (a) can be concrete and direct, or abstract, as in (b).

(a) have(L, Th) Mary has a book.

(b) have(L, Th) Mary has a problem with Bill.

In conclusion, we can give the following dependencies between thematic roles:



[5] Cf. Pustejovsky (1987) for an explanation of this term and a full discussion of the extended aspect calculus.

The generaliztion features apply to this structure to build hierarchical structures (Cf. [Keil 1979], [Kodratoff 1986]). This partial ordering allows us to define a notion of *covering*, as with a semi-lattice, from which a strong principle of functional uniqueness is derivable (cf. [Jackendoff 1985]). The mapping of a thematic role to an argument follows the following principle:

(9) **Maximal Assignment Principle** An argument will receive the maximal interpretation consistent with the data.

This says two things. First, it says that an *Agent*, for example, will always have a *location* and *theme* role associated with it. Furthermore, an *Agent* may be affected by its action, and hence be a *Patient* as well. Secondly, this principle says that although an argument may bear many thematic roles, the grammar picks out that function which is *maximally specific* in its interpretation, according to the markedness theory. Thus, the two arguments might be *Themes* in "John chased Mary", but the thematic roles which maximally characterize their functions in the sentence are A and P, respectively.

4. The Learning Component

4.1 The Form of the Input

The input is a data structure pair; an event sequence expression and a sentence describing the event. The event-sequence is a simulated output from a middle-level vision system where motion detection from the low-level input has already been associated with particular object types. [6]

The event-sequence consists of three instantaneous descriptions (IDs) of a situation represented as intervals. These correspond to the intervals t_1, t_2, and t_3 in the aspect calculus. The predicates are perceptual primitives, such as those described in Miller and Johnson-Laird (1976) and Maddox and Pustejovsky (1987), such as $[AT(t_1, e) \,\&\, e = [ON(a,c) \,\&\, Animate(a) \,\&\, Moves(a) \,\&\, \ldots]]$. The second object is a linguistic expression (i.e. a sentence), parsed by a simple finite state transducer. [7]

[6] For a detailed discussion of how the visual processing and linguistic systems interact, cf. Maddox and Pustejovsky (1987).

[7] We are not addressing any complex interaction between syntactic and semantic acquisition in this system. Ideally, we would like to integrate the concept acquisition mechanisms here with a parser such as Berwick's, Cf. Berwick 1985.

4.2 The Acquisition Procedure

We now turn to the design of the learning program itself. TULLY can be characterized as a domain-intensive inductive learning system, where the generalizations possible in the system are restricted by the architecture imposed by the semantic model. We can separate clearly what is given from what is learned in the system, as shown in Figure 1.

GIVEN
Extended Aspect Calculus
θ-Markedness Theory
Canonical Mapping
Rule Execution Loop

ACQUIRED
Verbal Lexical semantics
Argument-function mapping
Predication Hierarchy

Figure 1

In order to better understand the learning mechanism, we will step through an example run of the system. First, however, we will give the rule execution loop which the system follows.

Rule Execution Loop

1. *Instantiate Existing Thematic Indexes*
   INSTANTIATE: Attempt to do a semantic analysis of word given using existing Thematic Mapping Indexes. If the analysis fails then go to 2.

2. *Concept-acquisition phase.*
   Note failure: Credit assignment.
   Link arguments to roles according to Canonical Mapping.

3. *Build new Thematic Mapping Index*
   LINK and SHIFT: Constructs new index according to the Extended Aspect Calculus using information from credit assignment in (2). If this fails then go to (4).

4. *Invoke Noncanonical Mapping Principle.*
   If (3) fails to build a mapping for the lexical item in the input, then the rule INTERSECT is invoked. This allows the lines to cross from any of the interpretive levels to the argument tier.

5. *Generalization Step.*
   This is where the markedness theory is invoked. Induction follows the restrictions in the theory, where generalization is limited to one of the stated types.

Assume that the first input to the system is the sentence "Mary hit the cat," with its accompanying event sequence expression, represented as a situation calculus expression. INSTANTIATE attempts to map an existing *Thematic Mapping Index* onto the input, but fails. Stage (2) is entered by the failure of (1), and credit assignment indicates where it failed. Heuristics will indicate which thematic properties are associated with each argument, and stage (3) links the arguments with the proper roles, according to Canonical Mapping. This links *Mary* to *Agent* and *the cat* to *Patient*.

One important point to make here is that any information from the perceptual expression that is not grammatically expressed will automatically be assumed to be part of the verb meaning itself. In this case, the instrument of the hitting (e.g. Mary's arm) is covered by the lexical semantics of *hit*.

There are two forms of generalization performed by the system in step (5): *constraint propagation* and *thematic decoupling*. In a propagation procedure (Cf. [Waltz, 1975]), the computation is described as operating *locally*, since the change has local consistency. To illustrate, consider the verb entry for *have*, as in (1),

(1) John has a book. $have(x = L, y = Th)$

where the object carries the feature [$-abstract$]. Now, consider how the sense of the verb changes with a feature change to [$+abstract$], as in (2).

(2) John has an idea.

In other words, there is a propagation of this feature to the subject, where the sense of locative becomes more abstract, e.g. *mental*. These types of extensions give rise to other verbs with the same thematic mapping, but with "relaxed" interpretations. [8]

The other strategy employed here is that of *thematic decoupling*, where thematic information becomes disassociated from the lexical semantics for a verb. [9] The narrower interpretation of a verb's meaning will be arrived at after enough training instances are given; for example, from *cut* as meaning a particular action with a knife, to *cut* as an action that results in a certain state.

It is interesting to speculate on how these strategies facilitate the development from perceptual relations to more abstract ones. The verb *tell*, for example, can be viewed as a transfer verb with a [$+abstract$] *Theme*, and the accompanying contraint propagation (Cf. [Pinker, 1984] and [Jackendoff, 1983]). Similarly, experiencer verbs such as *please*, *upset*, and *anger* can be seen as combining both strategies: they are similar to transfer verbs, but with feature relaxation on the *Theme*, together with propagated constraints to the *Source* and *Goal* (the subject and object, respectively); the difference is that the *Theme* is incorporated and is not grammatically expressed.

John pleased his mother.

$please(x = S, y = G, Th: incorporated)$

Conclusions

In this paper we have outlined a theory of acquisition for the semantic roles associated with verbs. Specifically, we argue that perceptual predicates form the foundation for later conceptual development in language, and propose a specific algorithm for learning employing a theory of markedness for thematic types and the two strategies of *thematic decoupling* and *constraint relaxation and propagation*. The approach sketched above will doubtless need revision and refinement on particular points, but is claimed to offer a new perspective which can contribute to the solution of some long-standing puzzles in acquisition.

Acknowledgements

I would like to thank Sabine Bergler who did the first implementation of the algorithm, as well as Anthony Maddox, John Brolio, Ken Wexler, Mellissa Bowerman, and Edwin Williams for useful discussion. All faults and errors are of course my own.

References

[1] Angluin, D. "Inductive Inference of formal Languages from positive data." *Information and Control* 45:117-135.

[2] Berwick, Robert C. The Acquisition of Syntactic Information, MIT Press, Cambridge, MA. 1985.

[3] Berwick, Robert C., "Learning from Positive-Only Examples: The Subset Principle and Three Case Studies," in Michalski et al, 1986.

[4] Bowerman, Mellissa "Learning the Structure of Cau sative Verbs," in Clark (ed) *Papers and reports on child language development*, No. 8, Stanford University Committee on Linguistics. 1974

[5] Chomsky, Noam *Rules and Representation*, Columbia University Press, 1980

[6] Chomsky, Noam *Lectures on Government and Binding*, Foris, Holland, 1981.

[7] Dowty, David R., Word Meaning and Montague Grammar, D. Reidel, Dordrecht, Holland, 1979.

[8] Jackendoff, Ray, *Language and Cognition*, MIT Press, Cambridge, MA. 1983.

[9] Jackendoff, Ray, "The Role of Thematic Relations in Linguistic Theory,", ms. Brandeis University, 1985

[8] For further discussion of constraint propagation as a learning strategy, cf. Pustejovsky (1987b).

[9] Results given in Nygren (1977) indicate that children have fully incorporated instruments for verbs such as *hammer*, *cut*, and *saw*, and only at a later age do they abstract to a verb sense without a particular and constant instrument interpretation.

[10] Kodratoff, Yves, and J-G. Ganascia, "Improving the Generalization Step in Learning", in Michalskiet el (eds.), *Machine Learning II*, Morgan Kaufmann.

[11] Marcus, Mitch, *A Theory of Syntactic Recognition for Natural Language*, MIT Press, Cambridge, 1980

[12] Michalski, R.S., "A Theory and Methodology of Inductive Learning,", in Michalski et al (eds.), *Machine Learning I.*

[13] Miller, George, "Dictionaries of the Mind" in Proceedings of the 23rd Annual Meeting of the Association for Computational Linguistics, Chicago, 1985.

[14] Miller, George and Philip Johnson-Laird, Language and Perception, Belknap, Harvard University Press, Cambridge, MA. 1976.

[15] Mitchell, Tom, "Version Spaces: A Candidate Elimination Approach to Rule Learning," in IJCAI-77, 1977

[16] Mitchell, Tom, *Version Spaces: An Approach to Concept Learning*, Ph.D. thesis Stanford, 1978.

[17] Nygren, Carolyn, "Results of Experiments with Instrumentals," ms. UMASS, Amherst, MA.

[18] Pilato, Samuel F. and Robert C. Berwick, "Reversible Automata and Induction of the English Auxiliary System", in Proceedings of the 23rd Annual Meeting of the Association for Computational Linguistics, Chicago, 1985.

[19] Pinker, Steven, *Language Learnability and Language Development*, Harvard University Press, Cam bridge, 1984

[20] Pustejovsky, James, "A Theory of Lexical Semantics for Concept Acqusition in Natural Language", to appear in *International Journal of Intelligent Systems*

[21] Pustejovsky, James and Sabine Bergler, "On the Acquisition of the Conceptual Lexicon", paper submitted to AAAI-1987, Seattle, WA.

[22] Slobin , D. "Universals and Particulars in Language Acqusition", in Gleitmann, *Language Acquisition*, Cambridge, 1982

[23] Waltz, David "Understanding line drawings of scences with shadows," in *The Psychology of Computer Vision*, P. Winston ed. New York, McGraw-Hill, pp. 19-92.

[24] Waltz, David "Event Space Descriptions," Proceedings of the AAAI-82, 1982

[25] Williams, Edwin, "Predication", *Linguistic Inquiry*, 1980

[26] Winston, Patrick H., "Learning by Augmenting Rules and Accumulating Censors," in Michalski et al, 1986.

[27] Winston, Patrick, Binford, Katz, and Lowry, "Learn ing Physical Descriptions from Functional Definitions, Examples, and Precedents, Proceedings of AAAI, Washington, 1983