Title

Objects, actions, nouns, and verbs

Journal
Proceedings of the Annual Meeting of the Cognitive Science Society, 16(0)

Authors
Hastings, Peter M .
Lytinen, Steven L.

Publication Date
1994

Peer reviewed

Objects, actions, nouns, and verbs

Peter M. Hastings
Artificial Intelligence Lab
1101 Beal Avenue
The University of Michigan
Ann Arbor, MI 48109
(313)763-9074
peter@umich.edu

Steven L. Lytinen
DePaul University
Department of Computer Science
243 South Wabash Avenue
Chicago, IL 60604-2302
(312)362-6106
lytinen@cs.depaul.edu

Abstract

This paper describes a lexical acquisition mechanism that was implemented in order to increase the robustness of a Natural Language Processing system. Although the mechanism was not intended to be a cognitive model of children's language acquisition, it demonstrates many similarities with psycholinguistic findings. In particular, the structure of the domain knowledge representation forces the system to take a bipolar approach to learning nouns and verbs. Psycholinguistic studies demonstrate differing treatment of nouns and verbs by children and suggest a structural basis for this difference. The knowledge-level similarities between our system and human linguistic knowledge make it possible to infer that children must adopt a similar strategy to effectively learn word meanings.

Introduction

We have developed an incremental approach to learning word meanings from their usage in example sentences. This approach has been developed and implemented as part of the LINK Natural Language Processing system (Lytinen & Roberts 1989). The learning system, called Camille (Contextual Acquisition Mechanism for Incremental Lexeme Learning), selects candidate word meanings from a collection of frames, stored in a concept hierarchy which is part of LINK's knowledge base.

The original motivation for building Camille was to improve the robustness of LINK. With the addition of the learning mechanism, LINK can process sentences which contain words that are not already defined in the system's knowledge base. Although this original motivation did not include the desire to develop a cognitive model of human language acquisition, it has allowed some insight into the process of human lexical acquisition.[1] In particular, the structure of LINK's knowledge base, along with the nature of the task of acquiring new word meanings, forces Camille to take a different approach for learning nouns than for learning verbs. This dichotomy closely parallels psychological findings about human acquisition of nouns and verbs (Gentner 1978; Goldin-Meadow, Seligman, & Gelman 1976; Behrend 1990). Thus, Camille suggests an explanation for the differences described in the psycholinguistic literature.

Unlike other acquisition systems (Zernik 1987; Granger 1977; Salveter 1979; Selfridge 1986), and in accord with psycholinguistic theories like Syntactic Bootstrapping (Gleitman 1990; Naigles 1990), Camille takes full advantage of LINK's grammatical knowledge in making its inferences. The acquisition procedure takes place during parsing and interacts with the syntactic decisions made by the parser.

This paper describes the nature of the knowledge that Camille uses to make its inferences, the basic process,[2] and the correspondences to developmental psycholinguistics. The last section contains an analysis of the types of inferences that are affected by the structure of the knowledge representation.

The nature of the knowledge

LINK's knowledge representation consists of an inheritance hierarchy of domain-independent and domain-specific concepts, or frames. Figure 1 shows some of LINK's domain-specific object concepts which were used during our participation in ARPA's Third and Fourth Message Understanding Competitions, or MUC-3 and 4 (Sundheim 1992) (the shading will be explained later). The MUC task involved the extraction of information from newspaper articles which described terrorist activities in Latin America; thus, the frames encode knowledge about the domain of terrorism.

The contents of the knowledge base are determined by the requirements of the domain and by the need to cover the input language. The structure is determined by the relative specificity of the individual frames. The most general frames are located near the root of the hierarchy and specific ones are located at the leaves.

Figure 2 shows some of the actions from the terrorism domain. Action concepts provide the relational structure that binds together the representation of the meaning of sentences. These concepts also constrain the types of arguments that can be attached as their slot-fillers (the constraints are also shown in figure 2). The Detonate node, for example, has the constraint that its semantic object must be a descendant of Bomb (or the node Bomb itself). Detonate inherits an additional constraint from its parent Nasty-Action. This constraint specifies that the actor must be a Human.

The nodes in LINK's concept hierarchy serve as its basic units of meaning. For Camille, then, learning the meaning of an unknown word reduces to finding the appropriate node in the hierarchy which best represents the word's meaning — a graph search problem. To drive the search, the semantic constraints, which are normally used to limit attachment of slot-fillers to the head verb, interact with the "evidence" provided

[1]General correspondences of the initial implementation of this system with psycholinguistic findings were described in (Hastings, Lytinen, & Lindsay 1991).

[2]The actual process that Camille uses is described more fully in (Hastings 1994).

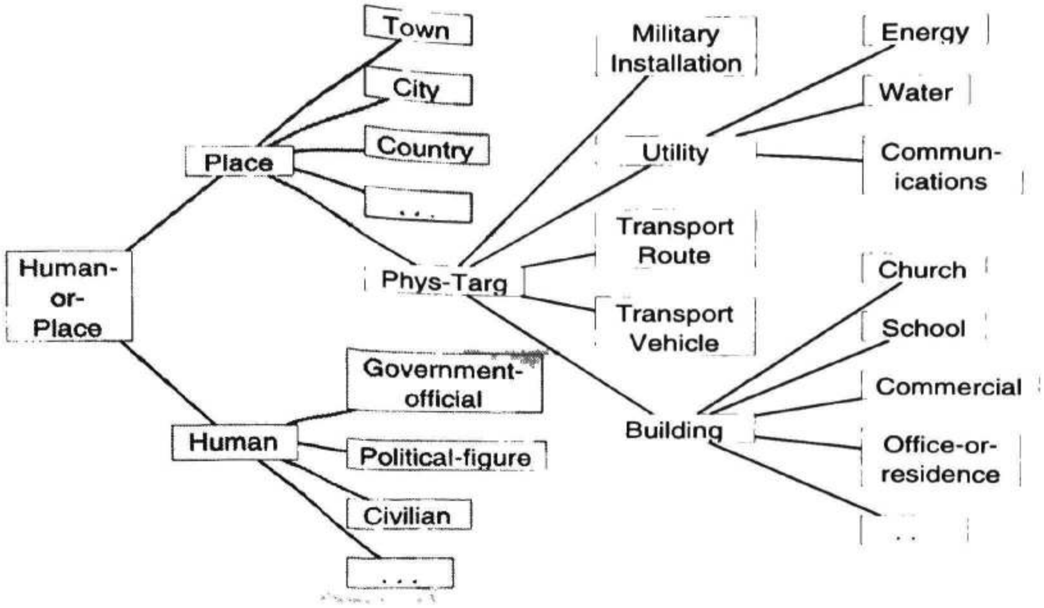

Figure 1: The pruned object tree

by example sentences. But the interaction works in different ways for different classes of words. Nouns (as the heads of noun phrases) normally serve as the slot-fillers of sentences and thus, as the items which are constrained. For example, in the sentence "Terrorists destroyed a flarge," the word "destroy" refers to the concept Destroy which has the constraint [Object = Phys-Targ]. When "flarge" is attached as the Object of the verb, the constraint places an *upper bound* on its interpretation as depicted in figure 1. The shaded-out nodes cannot be a valid interpretation of the meaning of "flarge".

For unknown verbs, however, the situation is quite different. Because they usually map to the actions in the domain, the verbs *apply* the constraints. Thus, the constraints place a *lower bound* on the interpretation of unknown verbs. The shaded areas of figure 2 show the concepts that are ruled out for an example sentence like "Terrorists froobled the headquarters." Possible meanings of "frooble" which are consistent with this sentence include any action which allows "headquarters" as its semantic object; thus, actions with constraints that are specific enough to exclude "headquarters" are not candidates for the meaning of "frooble."

Since input to Camille comes from the results of LINK's parses of example sentences from the domain, Camille does not receive negative examples as part of the learning process. Because of this, the learning problem is underconstrained for both nouns and verbs, in opposite ways. Example sentences provide an upper bound in the concept hierarchy on candidate hypotheses for nouns, but no lower bound, while the reverse is true for verbs. Because of this dichotomy, Camille must have different strategies for learning verbs and learning nouns. They can be stated most succinctly as follows:

- For nouns, choose the most general hypothesis which is consistent with the example sentence(s)
- For verbs, choose the most specific hypothesis which is consistent with the example(s)

In essence, because evidence only provides one kind of constraint (a lower or upper bound) for both nouns and verbs, Camille must form hypotheses for the meanings of unknown words which can be falsified by future examples. This means choosing the hypotheses which are closest to the bound provided by the evidence. More abstractly, then Camille's learning strategy could be stated as, "Always choose the concept(s) closest to the boundary." If this strategy were not pursued, future evidence would not allow Camille to further refine its hypotheses, thus resulting in overly general verb definitions and overly specific noun definitions.

It is important to note that the difference in information provided by example sentences for unknown nouns and verbs is is not just an artifact of LINK's knowledge representation structure. It is due to a fundamental principle of language. Because actions serve as the relational elements of sentence structure, they are the only logical place for the constraints to reside; thus contextual information about the meaning of a verb is intrinsically different from the information provided by context for unknown nouns. This difference is prescribed by the nature of the knowledge and it is consonant with psycholinguistic theories which will be described below. Without this two-part strategy, Camille could never learn those concepts closest to the edge because additional examples could never disconfirm the initial hypotheses. The next section briefly describes Camille's inference procedure.

Camille's inference procedure

As mentioned above, Camille takes a different approach to learning nouns than to learning verbs, but both start the same way. When Camille does not know the definition of a word, it enters a generic definition into the parse for various parts of speech. The combination of morphological and syntactic constraints is almost always sufficient to determine the correct lexical category for an unknown word. The rest of this section

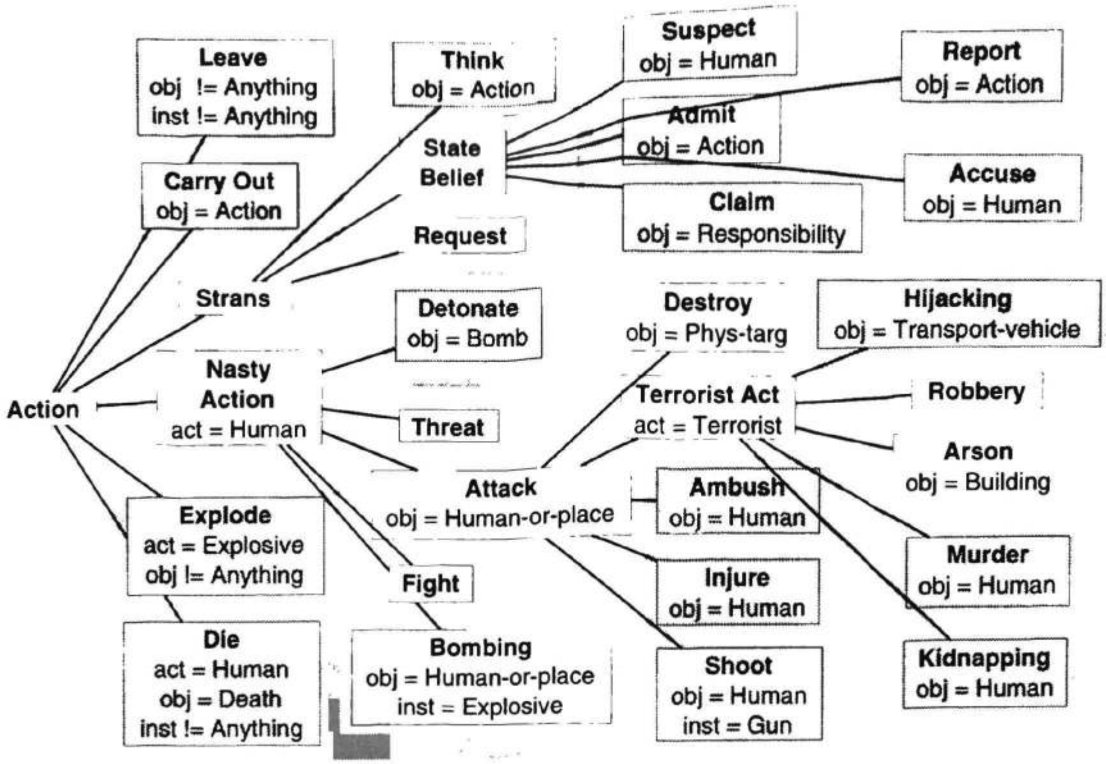

Figure 2: The pruned action tree

contrasts the learning of noun meanings and the learning of verb meanings.

Learning noun meanings

The process for learning nouns is fairly straightforward. When the noun is attached as an argument of a verb, any semantic constraints from that verb are applied to the noun. As mentioned above, this provides a single concept which is the root of a subtree that contains the valid interpretations for the word's meaning. Because the system wants to infer the most general noun meanings, this upper bound node is taken as the hypothesis for the word's meaning. Further examples can further specify the meaning.

As previously mentioned, given the sentence, "Terrorists destroyed a flarge," Camille can logically deduce that "flarge" must refer to some type of Phys-Targ. In practice, Camille infers that "flarge" means exactly the concept Phys-Targ. With a later sentence like, "Mary hijacked the flarge," Camille can further refine the hypothesis, inferring that "flarge" refers to Transport-Vehicle.

Learning verb meanings

For verbs, the process works the other way around. It is more difficult, however, because the system must search the action hierarchy for a node with appropriate constraints for the current slot fillers. This defines a lower bound on the set of possible meanings as described above. Unfortunately, the tree structure produces many concepts along the lower bound (8 nodes in figure 2). Camille ranks these hypotheses by computing the distance between their constraints and the actual slot filler in the example sentence. Camille keeps only the tightest fits, i.e. the concepts whose constraints most closely match the example. These concepts are the most specific valid hypotheses *and* the most falsifiable ones. Camille's learning method is incremental, so later examples of the word's use can either confirm the original hypothesis, or conflict with that hypothesis, triggering further search.

As an example of Camille's verb acquisition procedure, consider the sentence given above, "Terrorists froobled the headquarters." When LINK attaches "headquarters", which is represented by the Building concept, as the semantic object of the verb "frooble", the search for a meaning for "frooble" begins. As previously mentioned, the system can deductively rule out those concepts that were shaded out in figure 2. In order to *induce* a specific, falsifiable hypothesis, Camille prefers the concept Arson because its constraint that the object is a Building most closely matches the evidence.[3] The attachment of "Terrorists" as the actor of the verb is consistent with this hypothesis, so it is stored as the tentative definition for "frooble".

Given an additional example of the word's use like, "Joe froobled a pedestrian," Camille realizes that it has made an inappropriate guess because Arson's object constraint is violated by "pedestrian". Camille is incremental, so it takes the previous hypothesis as the starting point of the search for a better guess. The system climbs the tree (because any

[3]Because the object constraint for Arson is also the node Building, the distance between the constraint and the filler is 0. The object constraint for the Destroy node, on the other hand, is Phys-Targ which has a distance of 1 from Building.

more specific concept would also necessarily have constraints that were violated) and then tries to find another specific hypothesis that can take both a Civilian and a Building as its object. The node Robbery with its inherited object constraint Human-or-Place satisfies the search conditions and becomes the new hypothesis for the meaning of "frooble".

Empirical testing

The system was empirically evaluated on a set of 50 sentences chosen randomly from the MUC terrorism corpus.[4] The test sentences were quite long and complex, with an average length of 23 words, for example:

> Lopez Albujar, who left his post at the Ministry in May 1989, was riddled with bullets as he was getting out of his car in the Lima residential district of San Isidro.

To test Camille's learning algorithm, the definitions of all the verbs[5] in this set were removed from the lexicon. Then each of the sentences was processed by the system, and the learned definitions of these verbs were compared to the original definitions. Despite the complexity of the sentences and the low number of repetitions of each verb (average 2.7 occurrences), the system was able to infer the appropriate meanings for 8 of the 15 verbs, or 53%.[6]

Psychological connections

When psychologists first started studying the types of words that children acquired first, a striking observation was made. Children learn nouns well before and at a faster rate than they learn verbs. This prompted several studies of the differences in acquisition between the various types of words and theories about what causes those differences. These studies suggest that differences in the knowledge representation for nouns and verbs force children to use different mechanisms to learn them.

Gentner (1978, pp. 988-989) cites several studies that describe differences in acquisition between nouns and verbs. Some of these studies showed that young children's initial vocabulary consists entirely of nouns with verbs slowly making their way in. Others showed that the first verbs took almost twice as long to appear as the first nouns. A study of comprehension and production by Goldin-Meadow, Seligman, and Gelman (1976) showed two stages of early lexical development. In both stages, many more nouns than verbs were comprehended. Only a portion of the comprehended verbs were produced in the second stage, and none were produced in the first. Finally, Gentner described an additional study that demonstrated that the differences in acquisition are not just attributable to differences in the frequency of verbs versus nouns that the child hears. Even when presented with made-up nouns and verbs, and when balancing the presentation of these new words, children first used verbs an average of 8 months after starting to use nouns.[7]

In a slightly different vein, Behrend conducted in-depth studies of different types of verbs to compare children's comprehension and production among these various verbs (1990). The types of verbs that he studied were those that described actions (e.g. "squeeze", "pound"), results ("flatten", "break"), and instruments ("hammer"). He found that when labelling actions ("What is the person doing?"), children are more likely to use an instrument verb than an action verb. This seems strange for two reasons. First, because children use relatively little evidence and quickly infer mappings from words to meanings, they should make the best possible guess about the meaning of the word given what they know. All other things being equal, this should correspond to the type of verb that occurs most frequently in the language. But instrument verbs are far less frequent than action verbs are. Second, instrument verbs carry more information than action verbs and are therefore more specific. Thus the children in the experiments were labeling the events with the *most specific* label possible. This contradicts the results found in acquisition of nouns, which demonstrate that "specific subordinate terms are used much less frequently than basic-level terms as labels for familiar objects." (Behrend 1990, p. 694).

These results point to a fundamental difference between children's acquisition of nouns and verbs, one that can be most readily explained by differences in the underlying mental representation of the different concepts. Gentner calls the basis for this difference the "referential / relational" distinction. Nouns normally refer to objects or "thinglike elements." Objects tend to be highly constrained by the physical world. Hence, similar objects share almost all the same attributes. On the other hand, verbs tend to express relationships between objects or changes in those relationships. Relationships are more abstract and more difficult to describe with a set of attributes. The representation that Gentner espouses to represent these varying concepts is a semantic net, in which meanings are built up compositionally by referring to more basic elements of meaning.

Gentner's representation focuses on the representation of single nodes of meaning. Behrend supports Huttenlocher and Lui's proposal (1979) that these differences in behavior are caused by the overall structure of the representation. They suggest that objects are organized in a structured hierarchy so that nearby elements share many of the same features. The relational elements that are expressed by verbs are represented in a matrix structure with nodes connecting across the various object hierarchies. Graesser, Hopkinson, and Schmid (Graesser, Hopkinson, & Schmid 1987) have recently done

[4]The test set was kept relatively small to simulate a sparse-input learning task. The assumption is that most of the lexical definitions have been entered as part of the knowledge engineering of the system. Words that were overlooked in this process are not likely to be encountered frequently by the system, so the testing set contains a small number of examples of each word.

[5]The empirical testing focused on learning verb meanings because they were more difficult to acquire.

[6]The system was also tested in another domain in which the sentences were very simple (average length: 4). Because the parser produced many more correct parses, the system scored much better, 71%.

[7]These results are tempered somewhat by recent suggestions that the findings may be specific to the English language. Gopnik and Choi (1990), in a study of the correlation between linguistic and cognitive development, cite studies that Korean- and Japanese-speaking children show a higher use of verbs during the one-word stage than English-speaking children do. They attribute this difference in behavior to structural differences in the languages. Fernald and Morikawa (1993) found evidence that maternal speech to the children was responsible for the increased use of verbs by Japanese-speaking children.

experimental testing to support this hypothesis. The subjects were asked to sort sets of words by similarity. The findings suggested that people tend to sort nouns hierarchically while verbs were less structured and more "cross-classified." The cross-classification of Camille's concepts is discussed in the next section.

Camille's implications for learning

The first important point to make is that Camille's knowledge representation is not just an artifact of the particular system. The IS-A inheritance hierarchy is used widely in Artificial Intelligence, and for good reason. The structure facilitates generalization, an ability that is key to learning and reasoning. It also provides efficient storage of information. Furthermore, various psychological studies support the existence of hierarchical structures in the brain ((Kaplan, Weaver, & French 1990) and (Keil 1991), for example).

The use of a hierarchy to represent objects also seems intuitively correct. For representing actions and relations, the utility of this structure is not so obvious. As previously pointed out, although some psycholinguistic researchers have postulated a hierarchical scheme for their representation, recently the focus has turned to more "matrix-like" schemes. But the latter approach seems to conflict with the observation about the nature of constraints provided by the input and whether an upper bound or lower bound is created on the set of possible meanings for an unknown word. It is just not clear what "lower bound" would mean in a matrix-type organization.

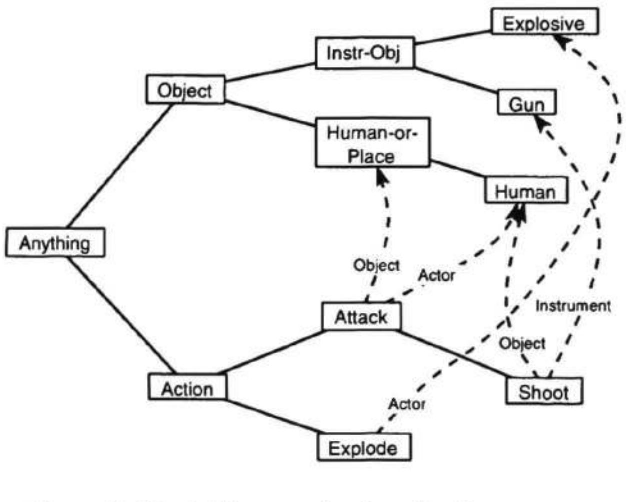

Figure 3: Matrix-like organization of action concepts

On close inspection, it appears that Camille's representation has the best of both worlds. If the slot-filler constraints are displayed graphically (see figure 3 where the solid lines represent paths in the IS-A hierarchy and the dashed lines represent constraints on the actions), it is apparent that the connections do (as Huttenlocher and Lui put it) cut across the various parts of the hierarchy. This leads to the question, "Does it make sense to have the additional structure imposed by enforcing a hierarchical structure on *actions*?" The answer appears to be yes, for the same reasons given for object representation above. The hierarchy has representational strength — it allows for efficient storage of the attributes and constraints of actions.

The question that remains then is, "What does this imply about human concept organization and the learning process?" For one thing, it lends support to the idea that there can be multiple organization structures within the brain. There are clear advantages to having both types of concept representation. In addition, it suggests that learning could proceed in one of two ways. If a child realizes that her idea of what a word means is wrong, she should look for concepts that are closely related in the hierarchy. If the child's hypothesized *constraints* for the word are wrong, she should change those constraints based on the structure of the hierarchy that is selected by the matrix links.

Conclusion

The structure of the concept knowledge that is used for inferring word meaning from context places constraints on how the learning process can proceed. Some form of these constraints should also apply to human lexical acquisition. The behavior of the computer model can therefore suggest explanations for psycholinguistic findings about differing treatment of nouns and verbs. It can also suggest other avenues of exploration to allow us to find out more about how children learn language.

Although Camille sheds some light on the difference in nature of learning nouns and verbs, in several ways the model is overly simple when compared to human lexical acquisition. First, the learning strategy as described here assumes that an unknown word has only one meaning. We have done preliminary work on extending the learning algorithm for ambiguous words. As described in (Hastings & Lytinen in press), the acquisition of ambigous words also has implications for the differing treatment of nouns and verbs. If the two-part strategy described here is not followed, the system cannot determine which words should have multiple definitions.

Second, the Camille model assumes the pre-existence of concepts which correspond to the meanings of new words. This is an overly simplistic assumption. It may be that learning a new word sometimes requires the addition of a new node in the concept hierarchy as well. Another extension of Camille (Hastings 1994) addresses concept acquisition within the context of a lexical acquisition system. In accordance with the results described herein, object concepts were found to be much easier to learn than action concepts.

Another limitation of the Camille system was brought to light by testing on larger data sets. Given more examples of the unknown words, the incremental system would be expected to increase the quality of its inferences. Instead the percentage of correct inferences decreased. This was due primarily to noise created by incorrect partial parses. As previously mentioned, the sentences in the terrorism domain were quite complex, and often the LINK parser could only return a partial parse tree. If the fragments included incorrect role assignments (e.g. "... Michigan froobled ..." instead of "... terrorists from Michigan froobled ..."), Camille misinterpreted this as justification for changing its previous hypothesis. Thus, the parser has a strong effect on the lexical acquisition process which must be further studied.

Even with these additional considerations, however, the essential differences between the tasks of learning nouns and verbs still remain in our model, and are intrinsic to the lexical acquisition task. This is due to the differences in usage of nouns and verbs. Thus, it seems inevitable that a computer system, or a human learning, must adopt different strategies when learning these two types of words.

References

Behrend, D. 1990. The development of verb concepts: Children's use of verbs to label familiar and novel events. *Child Development* 61:681–696.

Fernald, A., and Morikawa, H. 1993. Common themes and cultural variations in Japanese and American mothers' speech to infants. *Child Development* 64:637–656.

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

Gleitman, L. 1990. The structural sources of verb meaning. *Language Acquisition* I(1):3–55.

Goldin-Meadow, S.; Seligman, M.; and Gelman, R. 1976. Language in the two-year-old. *Cognition* 4:189–202.

Gopnik, A., and Choi, S. 1990. Do linguistic differences lead to cognitive differences? a cross-linguistic study of semantic and cognitive development. *First Language* 10:199–215.

Graesser, A.; Hopkinson, P.; and Schmid, C. 1987. Differences in interconcept organization between nouns and verbs. *Journal of Memory and Language* 26:242–253.

Granger, R. 1977. Foul-up: A program that figures out meanings of words from context. In *Proceedings of Fifth International Joint Conference on Artificial Intelligence*.

Hastings, P., and Lytinen, S. in press. The ups and downs of lexical acquisition. In *Proceedings of the* 12^{th} *National Conference on Artificial Intelligence*.

Hastings, P.; Lytinen, S.; and Lindsay, R. 1991. Learning words: Computers and kids. In Hammond, K., and Gentner, D., eds., *Proceedings of the 13th Annual Conference of the Cognitive Science Society*, 251–256. Hillsdale, NJ: Lawrence Erlbaum Associates.

Hastings, P. 1994. *Automatic Acquisition of Word Meaning from Context*. Ph.D. Dissertation, University of Michigan, Ann Arbor, MI.

Huttenlocher, J., and Lui, F. 1979. The semantic organization of some simple nouns and verbs. *Journal of verbal learning and verbal behavior* 18:141–162.

Kaplan, S.; Weaver, M.; and French, R. 1990. *Active symbols and internal models: Towards a cognitive connectionism*. London: Springer Verlag. Springer Series on Artificial Intelligence and Society.

Keil, F. 1991. Theories, concepts, and the acquisition of word meaning. In Byrnes, J. P., and Gelman, S. A., eds., *Perspectives on language and thought: Interrelations in development*. Cambridge: Cambridge University Press.

Lytinen, S., and Roberts, S. 1989. Unifying linguistic knowledge. AI Laboratory, Univ of Michigan, Ann Arbor, MI 48109.

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

Salveter, S. 1979. Inferring conceptual graphs. *Cognitive Science* 3:141–166.

Selfridge, M. 1986. A computer model of child language learning. *Artificial Intelligence* 29:171–216.

Sundheim, B. 1992. Overview of the fourth message understanding evaluation and conference. In *Proceedings of the Fourth Message Understanding Conference*. San Mateo, CA: Morgan Kaufmann Publishers, Inc.

Zernik, U. 1987. Strategies in language acquisitions: Learning phrases from examples in context. Technical Report UCLA-AI-87-1, UCLA.