To Appear in the Proceedings of the Fourteenth International Conference on Computational Linguistics, Nantes France, July 1992

Automatic Acquisition of Hyponyms from Large Text Corpora

Marti A. Hearst
Computer Science Division, 571 Evans Hall
University of California, Berkeley
Berkeley, CA 94720
and
Xerox Palo Alto Research Center
marti@cs.berkeley.edu

Abstract

We describe a method for the automatic acquisition of the hyponymy lexical relation from unrestricted text. Two goals motivate the approach: (i) avoidance of the need for pre-encoded knowledge and (ii) applicability across a wide range of text. We identify a set of lexico-syntactic patterns that are easily recognizable, that occur frequently and across text genre boundaries, and that indisputably indicate the lexical relation of interest. We describe a method for discovering these patterns and suggest that other lexical relations will also be acquirable in this way. A subset of the acquisition algorithm is implemented and the results are used to augment and critique the structure of a large hand-built thesaurus. Extensions and applications to areas such as information retrieval are suggested.

1 Introduction

Currently there is much interest in the automatic acquisition of lexical syntax and semantics, with the goal of building up large lexicons for natural language processing. Projects that center around extracting lexical information from Machine Readable Dictionaries (MRDs) have shown much success but are inherently limited, since the set of entries within a dictionary is fixed. In order to find terms and expressions that are not defined in MRDs we must turn to other textual resources. For this purpose, we view a text corpus not only as a source of information, but also as a source of information about the language it is written in.

When interpreting unrestricted, domain-independent text, it is difficult to determine in advance what kind of information will be encountered and how it will be expressed. Instead of interpreting everything in the text in great detail, we can search for specific lexical relations that are expressed in well-known ways. Surprisingly useful information can be found with only a very simple understanding of a text. Consider the following sentence:[1]

(S1) `The bow lute, such as the Bambara ndang, is plucked and has an individual curved neck for each string.`

Most fluent readers of English who have never before encountered the term *"Bambara ndang"* will nevertheless from this sentence infer that a *"Bambara ndang"* is a kind of *"bow lute"*. This is true even if the reader has only a fuzzy conception of what a bow lute is. Note that the author of the sentence is not deliberately defining the term, as would a dictionary or a children's book containing a didactic sentence like *A Bambara ndang is a kind of bow lute.* However, the semantics of the lexico-syntactic construction indicated by the pattern:

(1a) NP_0 *such as* $\{NP_1, NP_2 \ldots,$ *(and | or)*$\}$ NP_n

are such that they imply

(1b) *for all* NP_i, $1 \leq i \leq n$, *hyponym*(NP_i, NP_0)

Thus from sentence (S1) we conclude

hyponym("Bambara ndang", "bow lute").

We use the term *hyponym* similarly to the sense used in (Miller *et al.* 1990): a concept represented by a lexical item L_0 is said to be a hyponym of the concept represented by a lexical item L_1 if native speakers of English accept sentences constructed from the frame *An* L_0 *is a (kind of)* L_1. Here L_1 is the *hypernym* of L_0 and the relationship is reflexive and transitive, but not symmetric.

[1] All examples in this paper are real text, taken from *Grolier's American Academic Encyclopedia*.(Grolier 1990)

This example shows a way to discover a hyponymic lexical relationship between two or more noun phrases in a naturally-occurring text. This approach is similar in spirit to the pattern-based interpretation techniques being used in MRD processing. For example, (Alshawi 1987), in interpreting LDOCE definitions, uses a hierarchy of patterns which consist mainly of part-of-speech indicators and wildcard characters. (Markowitz *et al.* 1986), (Jensen & Binot 1987), and (Nakamura & Nagao 1988) also use pattern recognition to extract semantic relations such as taxonomy from various dictionaries. (Ahlswede & Evens 1988) compares an approach based on parsing Webster's 7th definitions with one based on pattern recognition, and finds that for finding simple semantic relations, pattern recognition is far more accurate and efficient than parsing. The general feeling is that the structure and function of MRDs makes their interpretation amenable to pattern-recognition techniques.

Thus one could say by interpreting sentence (S1) according to (1a-b) we are applying pattern-based relation recognition to general texts. Since one of the goals of building a lexical hierarchy automatically is to aid in the construction of a natural language processing program, this approach to acquisition is preferable to one that needs a complex parser and knowledge base. The tradeoff is that the the information acquired is coarse-grained.

There are many ways that the structure of a language can indicate the meanings of lexical items, but the difficulty lies in finding constructions that frequently and reliably indicate the relation of interest. It might seem that because free text is so varied in form and content (as compared with the somewhat regular structure of the dictionary) that it may not be possible to find such constructions. However, we have identified a set of lexico-syntactic patterns, including the one shown in (1a) above, that indicate the hyponymy relation and that satisfy the following desiderata:

(i) They occur frequently and in many text genres.
(ii) They (almost) always indicate the relation of interest.
(iii) They can be recognized with little or no pre-encoded knowledge.

Item (i) indicates that the pattern will result in the discovery of many instances of the relation, item (ii) that the information extracted will not be erroneous, and item (iii) that making use of the pattern does not require the tools that it is intended to help build.

Finding instances of the hyponymy relation is useful for several purposes:

Lexicon Augmentation. Hyponymy relations can be used to augment and verify existing lexicons, including ones built from MRDs. Section 3 of this paper describes an example, comparing results extracted from a text corpus with information stored in the noun hierarchy of WordNet ((Miller *et al.* 1990)), a hand-built lexical thesaurus.

Noun Phrase Semantics. Another purpose to which these relations can be applied is the identification of the general meaning of an unfamiliar noun phrases. For example, discovering the predicate

hyponym("broken bone", "injury")

indicates that the term "broken bone" can be understood at some level as an "injury" without having to determine the correct senses of the component words and how they combine. Note also that a term like "broken bone" is not likely to appear in a dictionary or lexicon, although it is a common locution.

Semantic Relatedness Information. There has recently been work in the detection of semantically related nouns via, for example, shared argument structures (Hindle 1990), and shared dictionary definition context (Wilks *et al.* 1990). These approaches attempt to infer relationships among lexical terms by looking at very large text samples and determining which ones are related in a statistically significant way. The technique introduced in this paper can be seen as having a similar goal but an entirely different approach, since only one sample need be found in order to determine a salient relationship (and that sample may be infrequently occurring or nonexistent).

Thinking of the relations discovered as closely related semantically instead of as hyponymic is most felicitous when the noun phrases involved are modified and atypical. Consider, for example, the predicate

hyponym("detonating explosive", "blasting agent").

This relation may not be a canonical ISA relation but the fact that it was found in a text implies that the terms' meanings are close. Connecting terms whose expressions are quite disparate but whose meanings are similar should be useful for improved synonym expansion in information retrieval and for finding chains of semantically related phrases, as used in the approach to recognition of topic boundaries of (Morris & Hirst 1991). We observe that terms that occur in a list are often related semantically, whether they occur in a hyponymy relation or not.

In the next section we outline a way to discover these lexico-syntactic patterns as well as illustrate those we have found. Section 3 shows the results of searching texts for a restricted version of one of the patterns and compares the results against a hand-built thesaurus. Section 4 is a discussion of the merits of this work and describes future directions.

2 Lexico-Syntactic Patterns for Hyponymy

Since only a subset of the possible instances of the hyponymy relation will appear in a particular form, we need to make use of as many patterns as possible. Below is a list of lexico-syntactic patterns that indicate the hyponymy relation, followed by illustrative sentence fragments and the predicates that can be derived from them (detail about the environment surrounding the patterns is omitted for simplicity):

(2) *such NP as* $\{NP\ ,\}^*$ $\{(or \mid and)\}$ *NP*
`... works by such authors as Herrick, Goldsmith, and Shakespeare.`
$\Longrightarrow$ *hyponym("author", "Herrick"), hyponym("author", "Goldsmith"), hyponym("author", "Shakespeare")*

(3) *NP* $\{,\ NP\}^*$ $\{,\}$ *or other NP*
`Bruises, wounds, broken bones or other injuries ...`
$\Longrightarrow$ *hyponym("bruise", "injury"), hyponym("wound", "injury"), hyponym("broken bone", "injury")*

(4) *NP* $\{,\ NP\}^*$ $\{,\}$ *and other NP*
`... temples, treasuries,and other important civic buildings.`
$\Longrightarrow$ *hyponym("temple", "civic building"), hyponym("treasury", "civic building")*

(5) *NP* $\{,\}$ *including* $\{NP\ ,\}^*$ $\{or \mid and\}$ *NP*
`All common-law countries, including Canada and England ...`
$\Longrightarrow$ *hyponym("Canada", "common-law country"), hyponym("England", "common-law country")*

(6) *NP* $\{,\}$ *especially* $\{NP\ ,\}^*$ $\{or \mid and\}$ *NP*
`... most European countries, especially France, England, and Spain.`
$\Longrightarrow$ *hyponym("France", "European country"), hyponym("England", "European country"), hyponym("Spain", "European country")*

When a relation *hyponym*(NP_0, NP_1) is discovered, aside from some lemmatizing and removal of unwanted modifiers, the noun phrase is left as an atomic unit, not broken down and analyzed. If a more detailed interpretation is desired, the results can be passed on to a more intelligent or specialized language analysis component. And, as mentioned above, this kind of discovery procedure can be a partial solution for a problem like noun phrase interpretation because at least part of the meaning of the phrase is indicated by the hyponymy relation.

2.1 Some Considerations

In example (4) above, the full noun phrase corresponding to the hypernym is *"other important civic buildings"*. This illustrates a difficulty that arises from using free text as the data source, as opposed to a dictionary – often the form that a noun phrase occurs in is not what we would like to record. For example, nouns frequently occur in their plural form and we usually want them to be singular. Adjectival quantifiers such as *"other"* and *"some"* are usually undesirable and can be eliminated in most cases without making the statement of the hyponym relation erroneous. Comparatives such as *"important"* and *"smaller"* are usually best removed, since their meaning is relative and dependent on the context in which they appear.

How much modification is desirable depends on the application to which the lexical relations will be put. For building up a basic, general-domain thesaurus, single-word nouns and very common compounds are most appropriate. For a more specialized domain, more modified terms have their place. For example, noun phrases in the medical domain often have several layers of modification which should be preserved in a taxonomy of medical terms.

Other difficulties and concerns are discussed in Section 3.

2.2 Discovery of New Patterns

How can these patterns be found? Initially we discovered patterns (1) - (3) by observation, looking through text and noticing the patterns and the relationships they indicate. In order to find new patterns automatically, we sketch the following procedure:

1. Decide on a lexical relation, R, that is of interest, e.g., *"group/member"* (in our formulation this is a subset of the hyponymy relation).
2. Gather a list of terms for which this relation is known to hold, e.g., *"England-country"*. This list can be found automatically using the method described here, bootstrapping from patterns found by hand, or by bootstrapping from an existing lexicon or knowledge base.
3. Find places in the corpus where these expressions occur syntactically near one another and record the environment.
4. Find the commonalities among these environments and hypothesize that common ones yield patterns that indicate the relation of interest.
5. Once a new pattern has been positively identified, use it to gather more instances of the target relation and go to Step 2.

We tried this procedure by hand using just one pair of terms at a time. In the first case we tried the *"England-country"* example, and with just this pair we found new patterns (4) and (5), as well as (1) - (3) which were already known. Next we tried *"tank-vehicle"* and discovered a very productive pattern, pattern (6). (Note that for this pattern, even though it has an emphatic element, this does not affect the fact that the relation indicated is hyponymic.)

We have tried applying this technique to meronymy (i.e., the part/whole relation), but without great success. The patterns found for this relation do not tend to uniquely identify it, but can be used to express other relations as well. It may be the case that in English the hyponymy relation is especially amenable to this kind of analysis, perhaps due to its "naming" nature. However, we have had some success at identification of more specific relations, such as patterns that indicate certain types of proper nouns.

We have not implemented an automatic version of this algorithm, primarily because Step 4 is underdetermined.

2.3 Related Work

This section discusses work in acquisition of lexical information from text corpora, although as mentioned earlier, significant work has been done in acquiring lexical information from MRDs.

(Coates-Stephens 1991) acquires semantic descriptions of proper nouns in a system called FUNES. FUNES attempts to fill in frame roles, (e.g., name, age, origin, position, and works-for, for a person frame) by processing newswire text. This system is similar to the work described here in that it recognizes some features of the context in which the proper noun occurs in order to identify some relevant semantic attributes. For instance, Coates-Stephens mentions that *"known as"* can explicitly introduce meanings for terms, as can appositives. We also have considered these markers, but the former often does not cleanly indicate "another name for" and the latter is difficult to recognize accurately. FUNES differs quite strongly from our approach in that, because it is able to fill in many kinds of frame roles, it requires a parser that produces a detailed structure, and it requires a domain-dependent knowlege base/lexicon.

(Velardi & Pazienza 1989) makes use of hand-coded selection restriction and conceptual relation rules in order to assign case roles to lexical items, and (Jacobs & Zernik 1988) uses extensive domain knowledge to fill in missing category information for unknown words.

Work on acquisition of syntactic information from text corpora includes Brent's (Brent 1991) verb subcategorization frame recognition technique and Smadja's (Smadja & McKeown 1990) collocation acquisition algorithm. (Calzolari & Bindi 1990) use corpus-based statistical association ratios to determine lexical information such as prepositional complementation relations, modification relations, and significant compounds.

Our methodology is similar to Brent's in its effort to distinguish clear pieces of evidence from ambiguous ones. The assumption is that that given a large enough corpus, the algorithm can afford wait until it encounters clear examples. Brent's algorithm relies on a clever trick: in the configuration of interest (in this case, verb valence descriptions), where noun phrases are the source of ambiguity, it uses only sentences which have pronouns in the crucial position, since pronouns do not allow this ambiguity. This approach is quite effective, but the disadvantage is that it isn't clear that it is applicable to any other tasks. The approach presented in this paper, using the algorithm sketched in the previous subsection, is potentially extensible.

3 Incorporating Results into WordNet

To validate this acquisition method, we compared the results of a restricted version of the algorithm with information found in WordNet.[2] WordNet (Miller *et al.* 1990) is a hand-built online thesaurus whose organization is modeled after the results of psycholinguistic research. To use the authors' words, Wordnet "... is an attempt to organize lexical information in terms of word meanings, rather than word forms. In that respect, WordNet resembles a thesaurus more than a dictionary ..." To this end, word forms with synonymous meanings are grouped into sets, called synsets. This allows a distinction to be made between senses of homographs. For example, the noun *"board"* appears in the synsets {*board, plank*} and {*board, committee*}, and this grouping serves for the most part as the word's definition. In version 1.1, WordNet contains about 34,000 noun word forms, including some compounds and proper nouns, organized into about 26,000 synsets. Noun synsets are organized hierarchically according to the hyponymy relation with implied inheritance and are further distinguished by values of features such as meronymy. WordNet's coverage and structure are impressive and provide a good basis for an automatic acquisition algorithm to build on.

When comparing a result *hyponym*(N_0,N_1) to the contents of WordNet's noun hierarchy, three kinds of outcomes are possible:

[2] The author thanks Miller, et al., for the distribution of WordNet.

Verify. If both N_0 and N_1 are in WordNet, and if the relation *hyponym*(N_0,N_1) is in the hierarchy (possibly through transitive closure) then the thesaurus is verified.

Critique. If both N_0 and N_1 are in WordNet, and if the relation *hyponym*(N_0,N_1) is *not* in the hierarchy (even through transitive closure) then the thesaurus is critiqued, i.e., a new set of hyponym connections is suggested.

Augment. If one or both of N_0 and N_1 are not present then these noun phrases and their relation are suggested as entries.

As an example of critiquing, consider the following sentence and derived relation:

(S2) `Other input-output devices, such as printers, color plotters, ...`
$\Longrightarrow$ *hyponym("printer", "input-output device")*

The text indicates that a printer is a kind of input-output device. Figure 1 indicates the portion of the hyponymy relation in WordNet's noun hierarchy that has to do with printers and devices. Note although the terms *device* and *printer* are present, they are not linked in such as way as to allow the easy insertion *I/O device* under the more general *device* and over the more specific *printer*. Although it is not obvious what to suggest to fix this portion of the hierarchy from this one relation alone, it is clear that its discovery highlights a trouble spot in the structure.

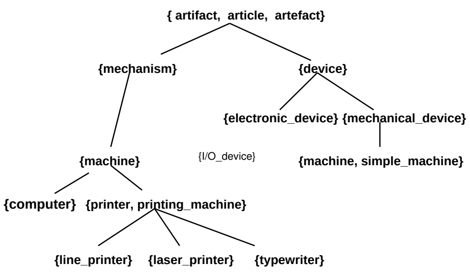

Figure 1: A Fragment of the WordNet Noun Hierarchy. Synsets are enclosed in braces; most synsets have more connections than those shown.

Most of the terms in WordNet's noun hierarchy are unmodified nouns or nouns with a single modifier. For this reason, in this experiment we only extracted relations consisting of unmodified nouns in both the hypernym and hyponym roles (although determiners are allowed and a very small set of quantifier adjectives: *"some", "many", "certain", and "other"*). Making this restriction is also useful because of the difficulties with determining which modifiers are significant, as touched on above, and because it seems easier to make a judgement call about the correctness of the classification of unmodified nouns for evaluation purposes.

Since we are trying to acquire lexical information our parsing mechanism should not be one that requires extensive lexical information. In order to detect the lexico-syntactic patterns, we use a unification-based constituent analyzer (taken from (Batali 1991)), which builds on the output of a part-of-speech tagger (Cutting *et al.* 1991). (All code described in this report is written in Common Lisp and run on Sun SparcStations.)

We wrote grammar rules for the constituent analyzer to recognize the pattern in (1a). As mentioned above, in this experiment we are detecting only unmodified nouns. Therefore, when a noun is found in the hypernym position, that is, before the lexemes *"such as"*, we check for the noun's inclusion in a relative clause, or as part of a larger noun phrase that includes an appositive or a parenthetical. Using the constituent analyzer, it is not necessary to parse the entire sentence; instead we look at just enough local context around the lexical items in the pattern to ensure that the nouns in the pattern are isolated.

After the hypernym is detected the hyponyms are identified. Often they occur in a list and each element in the list holds a hyponym relation with the hypernym. The main difficulty here lies in determining the extent of the last term in the list.

3.1 Results and Evaluation

Figure 2 illustrates some of the results of a run of the acquisition algorithm on *Grolier's American Academic Encyclopedia*(Grolier 1990), where a restricted version of pattern (1a) is the target (space constraints do not allow a full listing of the results). After the relations are found they are looked up in WordNet. We placed the WordNet noun hierarchy into a b-tree data structure for efficient retrieval and update and used a breadth-first-search to search through the transitive closure.

Out of 8.6M words of encyclopedia text, there are 7067 sentences that contain the lexemes *"such as"* contiguously. Out of these, 152 relations fit the restrictions of the experiment, namely that both the hyponyms and the hypernyms are unmodified (with the exceptions mentioned above). When the restrictions were eased slightly, so that NPs consisting of two nouns or a present/past participle plus a noun were allowed, 330 relations were found. When the latter experiment was run on about 20M words of *New York Times* text, 3178 sentences contained *"such as"* contiguously, and 46 relations were found using the strict no-modifiers criterion.

When the set of 152 *Grolier's* relations was looked up in WordNet, 180 out of the 226 unique words involved in the relations actually existed in the hierarchy, and

cereals:	rice* wheat*
countries:	Cuba Vietnam France*
hydrocarbon:	ethylene
substances:	bromine* hydrogen*
protozoa:	paramecium
liqueurs:	anisette* absinthe*
rocks:	granite*
substances:	phosphorus* nitrogen*
species:	steatornis oilbirds
bivalves:	scallop*
fungi:	smuts* rusts*
fabrics:	acrylics* nylon* silk*
antibiotics:	ampicillin erythromycin*
institutions:	temples king
seabirds:	penguins albatross*
flatworms:	tapeworms planaria
amphibians:	frogs*
waterfowl:	ducks
legumes:	lentils* beans* nuts
organisms:	horsetails ferns mosses
rivers:	Sevier Carson Humboldt
fruit:	olives* grapes*
hydrocarbons:	benzene gasoline
ideologies:	liberalism conservatism
industries:	steel iron shoes
minerals:	pyrite* galena
phenomena:	lightning*
infection:	meningitis
dyes:	quercitron

Figure 2: Relations found in *Grolier's*. The format is hypernym: hyponym list. Entries with * indicate relations found in WordNet.

61 out of the 106 feasible relations (i.e., relations in which both terms were already registered in WordNet) were found.

The quality of the relations found seems high overall, although there are difficulties. As to be expected, metonymy occurs, as seen in *hyponym("king", "institution")*. A more common problem is underspecification. For example, one relation is *hyponym("steatornis", "species")*, which is problematic because what *kind* of species needs to be known and most likely this information was mentioned in the previous sentence. Similarly, relations were found between *"device"* and *"plot", "metaphor", and "character"*, underspecifying the fact that literary devices of some sort are under discussion.

Sometimes the relationship expressed is slightly askance of the norm. For example, the algorithm finds *hyponym("Washington", "nationalist")* and *hyponym("aircraft", "target")* which are somewhat context and point-of-view dependent. This is not necessarily a problem; as mentioned above, finding alternative ways of stating similar notions is one of our goals. However, it is important to try to distinguish the more canonical and context-independent relations for entry in a thesaurus.

There are a few relations whose hypernyms are very high-level terms, e.g., *"substance"* and *"form"*. These are not incorrect; they just may not be as useful as more specific relations.

Overall, the results are encouraging. Although the number of relations found is small compared to the size of the text used, this situation can be greatly improved in several ways. Less stringent restrictions will increase the numbers, as the slight loosening shown in the *Grolier's* experiment indicates. A more savvy grammar for the constituent analyzer should also increase the results.

3.2 Automatic Updating

The question arises as to how to automatically insert relations between terms into the hierarchy. This involves two main difficulties. First, if both lexical expressions are present in the noun hierarchy but one or both have more than one sense, the algorithm must decide which senses to link together. We have preliminary ideas as to how to work around this problem. Say the hyponym in question has only one sense, but the hypernym has several. Then the task is simplified to determining which sense of the hypernym to link the hyponym to. We can then make use of a lexical disambiguation algorithm, e.g., (Hearst 1991), to determine which sense of the hypernym is being used in the sample sentence.

Furthermore, since we've assumed the hyponym has only one main sense we could do the following: Look through a corpus for occurrences of the hyponym and see if its environment tends to be similar to one of the senses of its hypernym. For example, if the hypernym is *"bank"* and the hyponym is *"First National"*, every time, within a sample of text, the term *"First National"* occurs, replace it with *"bank"*, and then run the disambiguation algorithm as usual. If this term can be positively classified as having one sense of bank over the others, then this would provide strong evidence as to which sense of the hypernym to link the hyponym to. This idea is purely speculative; we have not yet tested it.

The second main problem with inserting new relations arises when one or both terms do not occur in the hierarchy at all. In this case, we have to determine which, if any, existing synset the term belongs in and then do the sense determination mentioned above.

4 Conclusions

We have described a low-cost approach for automatic acquisition of semantic lexical relations from unrestricted text. This method is meant to provide an incremental step toward the larger goals of natural language processing. Our approach is complementary to statistically based approaches that find semantic relations between terms, in that ours requires a single specially expressed instance of a relation while the others require a statistically significant number of generally expressed relations. We've shown that our approach is also useful as a critiquing component for existing knowledge bases and lexicons.

We plan to test the pattern discovery algorithm on more relations and on languages other than English (depending on the corpora available). We would also like to do some analysis of the noun phrases that are acquired, and to explore the effects of various kinds of modifiers on the appropriateness of the noun phrase. We plan to do this in the context of analyzing environmental impact reports.

Acknowledgements. This work was supported in part by an internship at the Xerox Palo Alto Research Center and in part by the University of California and Digital Equipment Corporation under Digital's flagship research project Sequoia 2000: Large Capacity Object Servers to Support Global Change Research.

References

Ahlswede, T. & M. Evens (1988). Parsing vs. text processing in the analysis of dictionary definitions. *Proceedings of the 26th Annual Meeting of the Association for Computational Linguistics*, pages 217–224.

Alshawi, H. (1987). Processing dictionary definitions with phrasal pattern hierarchies. *American Journal of Computational Linguistics*, 13(3):195–202.

Batali, J. (1991). *Automatic Acquisition and Use of Some of the Knowledge in Physics Texts.* PhD thesis, Massachusetts Institute of Technology, Artificial Intelligence Laboratory.

Brent, M. R. (1991). Automatic acquisition of subcategorization frames from untagged, free-text corpora. In *Proceedings of the 29th Annual Meeting of the Association for Computational Linguistics.*

Calzolari, N. & R. Bindi (1990). Acquisition of lexical information from a large textual italian corpus. In *Proceedings of the Thirteenth International Conference on Computational Linguistics*, Helsinki.

Coates-Stephens, S. (1991). Coping with lexical inadequacy – the automatic acquisition of proper nouns from news text. In *The Proceedings of the 7th Annual Conference of the UW Centre for the New OED and Text Research: Using Corpora*, pages 154–169, Oxford.

Cutting, D., J. Kupiec, J. Pedersen, & P. Sibun (1991). A practical part-of-speech tagger. Submitted to *The 3rd Conference on Applied Natural Language Processing.*

Grolier (1990). *Academic American Encyclopedia.* Grolier Electronic Publishing, Danbury, Connecticut.

Hearst, M. A. (1991). Noun homograph disambiguation using local context in large text corpora. In *The Proceedings of the 7th Annual Conference of the UW Centre for the New OED and Text Research: Using Corpora*, pages 1–22, Oxford.

Hindle, D. (1990). Noun classification from predicate-argument structures. *Proceedings of the 28th Annual Meeting of the Association for Computational Linguistics*, pages 268–275.

Jacobs, P. & U. Zernik (1988). Acquiring lexical knowledge from text: A case study. In *Proceedings of AAAI88*, pages 739–744.

Jensen, K. & J.-L. Binot (1987). Disambiguating prepositional phrase attachments by using on-line dictionary definitions. *American Journal of Computational Linguistics*, 13(3):251–260.

Markowitz, J., T. Ahlswede, & M. Evens (1986). Semantically significant patterns in dictionary definitions. *Proceedings of the 24th Annual Meeting* *of the Association for Computational Linguistics*, pages 112–119.

Miller, G. A., R. Beckwith, C. Fellbaum, D. Gross, & K. J. Miller (1990). Introduction to WordNet: An on-line lexical database. *Journal of Lexicography*, 3(4):235–244.

Morris, J. & G. Hirst (1991). Lexical cohesion computed by thesaural relations as an indicator of the structure of text. *Computational Linguistics*, 17(1):21–48.

Nakamura, J. & M. Nagao (1988). Extraction of semantic information from an ordinary english dictionary and its evaluation. In *Proceedings of the Twelfth International Conference on Computational Linguistics*, pages 459–464, Budapest.

Smadja, F. A. & K. R. McKeown (1990). Automatically extracting and representing collocations for language generation. In *Proceedings of the 28th Annual Meeting of the Association for Computational Linguistics*, pages 252–259.

Velardi, P. & M. T. Pazienza (1989). Computer aided interpretation of lexical cooccurrences. *Proceedings of the 27th Annual Meeting of the Association for Computational Linguistics*, pages 185–192.

Wilks, Y. A., D. C. Fass, C. ming Guo, J. E. McDonald, T. Plate, & B. M. Slator (1990). Providing machine tractable dictionary tools. *Journal of Machine Translation*, 2.