CORPUS-BASED ACQUISITION OF RELATIVE PRONOUN DISAMBIGUATION HEURISTICS

Claire Cardie
Department of Computer Science
University of Massachusetts
Amherst, MA 01003
E-mail: cardie@cs.umass.edu

ABSTRACT

This paper presents a corpus-based approach for deriving heuristics to locate the antecedents of relative pronouns. The technique duplicates the performance of hand-coded rules and requires human intervention only during the training phase. Because the training instances are built on parser output rather than word cooccurrences, the technique requires a small number of training examples and can be used on small to medium-sized corpora. Our initial results suggest that the approach may provide a general method for the automated acquisition of a variety of disambiguation heuristics for natural language systems, especially for problems that require the assimilation of syntactic and semantic knowledge.

1 INTRODUCTION

State-of-the-art natural language processing (NLP) systems typically rely on heuristics to resolve many classes of ambiguities, e.g., prepositional phrase attachment, part of speech disambiguation, word sense disambiguation, conjunction, pronoun resolution, and concept activation. However, the manual encoding of these heuristics, either as part of a formal grammar or as a set of disambiguation rules, is difficult because successful heuristics demand the assimilation of complex syntactic and semantic knowledge. Consider, for example, the problem of prepositional phrase attachment. A number of purely structural solutions have been proposed including the theories of Minimal Attachment (Frazier, 1978) and Right Association (Kimball, 1973). While these models may suggest the existence of strong syntactic preferences in effect during sentence understanding, other studies provide clear evidence that purely syntactic heuristics for prepositional phrase attachment will not work (see (Whittemore, Ferrara, & Brunner, 1990), (Taraban, & McClelland, 1988)). However, computational linguists have found the manual encoding of disambiguation rules — especially those that merge syntactic and semantic constraints — to be difficult, time-consuming, and prone to error. In addition, hand-coded heuristics are often incomplete and perform poorly in new domains comprised of specialized vocabularies or a different genre of text.

In this paper, we focus on a single ambiguity in sentence processing: locating the antecedents of relative pronouns. We present an implemented corpus-based approach for the automatic acquisition of disambiguation heuristics for that task. The technique uses an existing hierarchical clustering system to determine the antecedent of a relative pronoun given a description of the clause that precedes it and requires only minimal syntactic parsing capabilities and a very general semantic feature set for describing nouns. Unlike other corpus-based techniques, only a small number of training examples is needed, making the approach practical even for small to medium-sized on-line corpora. For the task of relative pronoun disambiguation, the automated approach duplicates the performance of hand-coded rules and makes it possible to compile heuristics tuned to a new corpus with little human intervention. Moreover, we believe that the technique may provide a general approach for the automated acquisition of disambiguation heuristics for additional problems in natural language processing.

In the next section, we briefly describe the task of relative pronoun disambiguation. Sections 3 and 4 give the details of the acquisition algorithm and evaluate its performance. Problems with the approach and extensions required for use with large corpora of unrestricted text are discussed in Section 5.

2 DISAMBIGUATING RELATIVE PRONOUNS

Accurate disambiguation of relative pronouns is important for any natural language processing system that hopes to process real world texts. It is especially a concern for corpora where the sentences tend to be long and information-packed. Unfortunately, to understand a sentence containing a relative pronoun, an NLP system must solve two difficult problems: the system has to locate the antecedent of the relative pronoun and then determine the antecedent's implicit position in the embedded clause. Although finding the gap in the embedded clause is an equally difficult problem, the work we describe here focuses on locating the relative pronoun antecedent.[1]

This task may at first seem relatively simple: the antecedent of a relative pronoun is just the most recent constituent that is a human. This is the case for sentences S1-S7 in Figure 1, for example. However, this strategy assumes that the NLP system produces a perfect syntactic and semantic parse of the clause preceding the relative pronoun, including prepositional phrase attachment (e.g., S3, S4, and S7) and interpretation of conjunctions (e.g., S4, S5, and S6) and appositives (e.g., S6). In S5, for example, the antecedent is the entire conjunction of phrases (i.e., "Jim, Terry, and Shawn"), not just the most recent human (i.e., "Shawn"). In S6, either

S1. Tony saw *the boy* who won the award.
S2. *The boy* who gave me the book had red hair.
S3. Tony ate dinner with *the men* from Detroit who sold computers.
S4. I spoke to *the woman* with the black shirt and ☐ green hat over in the far corner of the room wh wanted a second interview.
S5. I'd like to thank *Jim, Terry, and Shawn*, who provided the desserts.
S6. I'd like to thank *our sponsors, GE and NSF*, who provide financial support.
S7. *The woman* from Philadelphia who played soccer was my sister.
S8. The awards for *the children* who pass the test are in the drawer.
S9. We wondered who stole the watch.
S10. We talked with *the woman and the man* who danced.

***Figure 1*. Examples of Relative Pronoun Antecedents**

"our sponsors" or its appositive "GE and NSF" is a semantically valid antecedent. Because pp-attachment and interpretation of conjunctions and appositives remain difficult for current systems, it is often unreasonable to expect reliable parser output for clauses containing those constructs.

Moreover, the parser must access both syntactic and semantic knowledge in finding the antecedent of a relative pronoun. The syntactic structure of the clause preceding "who" in S7 and S8, for example, is identical (NP-PP) but the antecedent in each case is different. In S7, the antecedent is the subject, "the woman;" in S9, it is the prepositional phrase modifier, "the children." Even if we assume a perfect parse, there can be additional complications. In some cases the antecedent is not the most recent constituent, but is a modifier of that constituent (e.g., S8). Sometimes there is no apparent antecedent at all (e.g., S9). Other times the antecedent is truly ambiguous without seeing more of the surrounding context (e.g., S10).

As a direct result of these difficulties, NLP system builders have found the manual coding of rules that find relative pronoun antecedents to be very hard. In addition, the resulting heuristics are prone to errors of omission and may not generalize to new contexts. For example, the UMass/MUC-3 system[2] began with 19 rules for finding the antecedents of relative pronouns. These rules included both structural and semantic knowledge and were based on approximately 50 instances of relative pronouns. As counter-examples were identified, new rules were added (approximately 10) and existing rules changed. Over time, however, we became increasingly reluctant to modify the rule set because the global effects of local rule changes were difficult to measure. Moreover, the original rules were based on sentences that UMass/MUC-3 had found to contain important information. As a result, the rules tended to work well for relative pronoun disambiguation in sentences of this class (93% correct for one test set of 50 texts), but did not generalize to sentences outside of the class (78% correct on the same test set of 50 texts).

2.1 CURRENT APPROACHES

Although descriptions of NLP systems do not usually include the algorithms used to find relative pronoun antecedents, current high-coverage parsers seem to employ one of 3 approaches for relative pronoun disambiguation. Systems that use a formal syntactic grammar often directly encode information for relative pronoun disambiguation in the grammar. Alternatively, a syntactic filter is applied to the parse tree and any noun phrases for which coreference with the relative pronoun is syntactically legal (or, in some cases, illegal) are passed to a semantic component which determines the antecedent using inference or preference rules (see (Correa, 1988), (Hobbs, 1986), (Ingria, & Stallard, 1989), (Lappin, & McCord, 1990)). The third approach employs hand-coded disambiguation heuristics that rely mainly on

---

[1]For a solution to the gap-finding problem that is consistent with the simplified parsing strategy presented below, see (Cardie & Lehnert, 1991).

[2]UMass/MUC-3 is a version of the CIRCUS parser (Lehnert, 1990) developed for the MUC-3 performance evaluation. See (Lehnert et. al., 1991) for a description of UMass/MUC-3. MUC-3 is the Third Message Understanding System Evaluation and Message Understanding Conference (Sundheim, 1991).

semantic knowledge but also include syntactic constraints (e.g., UMass/MUC-3).

However, there are problems with all 3 approaches in that 1) the grammar must be designed to find relative pronoun antecedents for all possible syntactic contexts; 2) the grammar and/or inference rules require tuning for new corpora; and 3) in most cases, the approach unreasonably assumes a completely correct parse of the clause preceding the relative pronoun. In the remainder of the paper, we present an automated approach for deriving relative pronoun disambiguation rules. This approach avoids the problems associated with the manual encoding of heuristics and grammars and automatically tailors the disambiguation decisions to the syntactic and semantic profile of the corpus. Moreover, the technique requires only a very simple parser because input to the clustering system that creates the disambiguation heuristics presumes neither pp-attachment nor interpretation of conjunctions and appositives.

3 AN AUTOMATED APPROACH

Our method for deriving relative pronoun disambiguation heuristics consists of the following steps:

1. Select from a subset of the corpus all sentences containing a particular relative pronoun. (For the remainder of the paper, we will focus on the relative pronoun "who.")

2. For each instance of the relative pronoun in the selected sentences,
   a. parse the portion of the sentence that precedes it into low-level syntactic constituents
   b. use the results of the parse to create a training instance that represents the disambiguation decision for this occurrence of the relative pronoun.

3. Provide the training instances as input to an existing conceptual clustering system.

During the training phase outlined above, the clustering system creates a hierarchy of relative pronoun disambiguation decisions that replace the hand-coded heuristics. Then, for each new occurrence of the wh-word encountered after training, we retrieve the most similar disambiguation decision from the hierarchy using a representation of the clause preceding the wh-word as the probe. Finally, the antecedent of the retrieved decision guides the selection of the antecedent for the new occurrence of the relative pronoun. Each step of the training and testing phases will be explained further in the sections that follow.

3.1 SELECTING SENTENCES FROM THE CORPUS

For the relative pronoun disambiguation task, we used the MUC-3 corpus of 1500 articles that range from a single paragraph to over one page in length. In theory, each article describes one or more terrorist incidents in Latin America. In practice, however, about half of the texts are actually irrelevant to the MUC task. The MUC-3 articles consist of a variety of text types including newspaper articles, TV news reports, radio broadcasts, rebel communiques, speeches, and interviews. The corpus is relatively small – it contains approximately 450,000 words and 18,750 sentences. In comparison, most corpus-based algorithms employ substantially larger corpora (e.g., 1 million words (de Marcken, 1990), 2.5 million words (Brent, 1991), 6 million words (Hindle, 1990), 13 million words (Hindle, & Rooth, 1991)).

Relative pronoun processing is especially important for the MUC-3 corpus because approximately 25% of the sentences contain at least one relative pronoun.[3] In fact, the relative pronoun "who" occurs in approximately 1 out of every 10 sentences. In the experiment described below, we use 100 texts containing 176 instances of the relative pronoun "who" for training. To extract sentences containing a specific relative pronoun, we simply search the selected articles for instances of the relative pronoun and use a preprocessor to locate sentence boundaries.

3.2 PARSING REQUIREMENTS

Next, UMass/MUC-3 parses each of the selected sentences. Whenever the relative pronoun "who" is recognized, the syntactic analyzer returns a list of the low-level constituents of the preceding clause prior to any attachment decisions (see Figure 2). UMass/MUC-3 has a simple, deterministic, stack-oriented syntactic analyzer based on the McEli parser (Schank, & Riesbeck, 1981). It employs lexically-indexed local syntactic knowledge to segment incoming text into noun phrases, prepositional phrases, and verb phrases, ignoring all unexpected constructs and unknown words.[4] Each constituent

---

[3]There are 4707 occurrences of wh-words (i.e., who, whom, which, whose, where, when, why) in the approximately 18,750 sentences that comprise the MUC-3 corpus.

[4]Although UMass/MUC-3 can recognize other syntactic classes, only noun phrases, prepositional phrases, and verb phrases become part of the training instance.

Sources in downtown Lima report that the police last night detained Juan Bautista and Rogoberto Matute, who ...

↓ **UMass/MUC-3 syntactic analyzer**

the police : *[subject, human]*
detained : *[verb]*
Juan Bautista : *[np, proper-name]*
Rogoberto Matute : *[np, proper-name]*

***Figure 2*. Syntactic Analyzer Output**

returned by the parser (except the verb) is tagged with the semantic classification that best describes the phrase's head noun. For the MUC-3 corpus, we use a set of 7 semantic features to categorize each noun in the lexicon: human, proper-name, location, entity, physical-target, organization, and weapon. In addition, clause boundaries are detected using a method described in (Cardie, & Lehnert, 1991).

It should be noted that all difficult parsing decisions are delayed for subsequent processing components. For the task of relative pronoun disambiguation, this means that the conceptual clustering system, not the parser, is responsible for recognizing all phrases that comprise a conjunction of antecedents and for specifying at least one of the semantically valid antecedents in the case of appositives. In addition, pp-attachment is more easily postponed until after the relative pronoun antecedent has been located. Consider the sentence "I ate with the men from the restaurant in the club." Depending on the context, "in the club" modifies either "ate" or "the restaurant." If we know that "the men" is the antecedent of a relative pronoun, however (e.g., "I ate with *the men* from the restaurant in the club, *who* offered me the job"), it is probably the case that "in the club" modifies "the men."

Finally, because the MUC-3 domain is sufficiently narrow in scope, lexical disambiguation problems are infrequent. Given this rather simplistic view of syntax, we have found that a small set of syntactic predictions covers the wide variety of constructs in the MUC-3 corpus.

3.3 CREATING THE TRAINING INSTANCES

Output from the syntactic analyzer is used to generate a training instance for each occurrence of the relative pronoun in the selected sentences. A training instance represents a single disambiguation decision and includes one *attribute-value pair* for every low-level syntactic constituent in the preceding clause. The attributes of a training instance describe the syntactic class of the constituent as well as its position with respect to the relative pronoun. The value associated with an attribute is the semantic feature of the phrase's head noun. (For verb phrases, we currently note only their presence or absence using the values *t* and *nil,* respectively.)

Consider the training instances in Figure 3. In S1, for example, "of the 76th district court" is represented with the attribute *pp1* because it is a prepositional phrase and is in the first position to the left of "who." Its value is "physical-target" because "court" is classified as a physical-target in the lexicon. The subject and verb constituents (e.g., "her DAS bodyguard" in S3 and "detained" in S2) retain their traditional *s* and *v* labels, however — no positional information is included for those attributes.

S1: [The judge] [of the 76th court] [,] who ...

Training instance: *[(s human) (pp1 physical-target) (v nil) (antecedent ((s)))]*

S2: [The police] [detained] [Juan Bautista] [and] [Rogoberto Matute] [,] who ...

Training instance: *[(s human) (v t) (np2 proper-name) (np1 proper-name) (antecedent ((np2 np1)))]*

S3: [Her DAS bodyguard] [,] [Dagoberto Rodriquez] [,] who...

Training instance: *[(s human) (np1 proper-name) (v nil) (antecedent ((np1)(s np1)(s)))]*

***Figure 3* . Training Instances**

In addition to the *constituent* attribute-value pairs, a training instance contains an attribute-value pair that represents the correct antecedent. As shown in Figure 3, the value of the *antecedent* attribute is a list of the syntactic constituents that contain the antecedent (or *(none)* if the relative pronoun has no antecedent). In S1, for example, the antecedent of "who" is "the judge." Because this phrase is located in the subject position, the value of the antecedent attribute is *(s)*. Sometimes, however, the antecedent is actually a conjunction of phrases. In these cases, we represent the antecedent as a list of the constituents associated with each element of the conjunction. Look, for example, at the antecedent in S2. Because "who" refers to the conjunction "Juan Bautista and Rogoberto Matute," and because those phrases occur as *np1* and *np2*, the value of the antecedent attribute is *(np2 np1)*. S3 shows yet another variation of the antecedent attribute-value pair. In this example, an appositive creates three equivalent antecedents: 1) "Dagoberto Rodriguez" —*(np1)*, 2) "her DAS bodyguard" — *(s)*, and 3) "her DAS bodyguard, Dagoberto Rodriguez" — *(s np1)*.

UMass/MUC-3 automatically generates the training instances as a side effect of parsing. Only the desired antecedent is specified by a human supervisor via a menu-driven interface that displays the antecedent options.

3.4 BUILDING THE HIERARCHY OF DISAMBIGUATION HEURISTICS

As the training instances become available they are input to an existing conceptual clustering system called COBWEB (Fisher, 1987).[5] COBWEB employs an evaluation metric called *category utility* (Gluck, & Corter, 1985) to incrementally discover a classification hierarchy that covers the training instances.[6] It is this hierarchy that replaces the hand-coded disambiguation heuristics. While the details of COBWEB are not necessary, it is important to know that nodes in the hierarchy represent concepts that increase in generality as they approach the root of the tree. Given a new instance to classify, COBWEB retrieves the most specific concept that adequately describes the instance.

3.5 USING THE DISAMBIGUATION HEURISTICS HIERARCHY

After training, the resulting hierarchy of relative pronoun disambiguation decisions supplies the antecedent of the wh-word in new contexts. Given a novel sentence containing "who," UMass/MUC-3 generates a set of attribute-value pairs that represent the clause preceding the wh-word. This probe is just a training instance without the antecedent attribute-value pair. Given the probe, COBWEB retrieves from the hierarchy the individual instance or abstract class that is most similar and the antecedent of the retrieved example guides selection of the antecedent for the novel case. We currently use the following selection heuristics to 1) choose an antecedent for the novel sentence that is consistent with the context of the probe; or to 2) modify the retrieved antecedent so that it is applicable in the current context:

1. Choose the first option whose constituents are all present in the probe.

2. Otherwise, choose the first option that contains at least one constituent present in the probe and ignore those constituents in the retrieved antecedent that are missing from the probe.

3. Otherwise, replace the np constituents in the retrieved antecedent that are missing from the probe with pp constituents (and vice versa), and try 1 and 2 again.

In S1 of Figure 4, for example, the first selection heuristic applies. The retrieved instance specifies the *np2* constituent as the location of the antecedent and the probe has *np2* as one of its constituents. Therefore, UMass/MUC-3 infers that the antecedent of "who" for the current sentence is "the hardliners," i.e., the contents of the *np2* syntactic constituent. In S2, however, the retrieved concept specifies an antecedent from five constituents, only two of which are actually present in the probe. Therefore, we ignore the missing constituents *pp5*, *np4*, and *pp3*, and look to just *np2* and *np1* for the antecedent. For S3, selection heuristics 1 and 2 fail because the probe contains no *pp2* constituent. However, if we replace *pp2* with *np2* in the retrieved antecedent, then heuristic 1 applies and "a specialist" is chosen as the antecedent.

[5] For these experiments, we used a version of COBWEB developed by Robert Williams at the University of Massachusetts at Amherst.

[6] Conceptual clustering systems typically discover appropriate classes as well as the the concepts for each class when given a set of examples that have **not** been preclassified by a teacher. Our unorthodox use of COBWEB to perform supervised learning is prompted by plans to use the resulting hierarchy for tasks other than relative pronoun disambiguation.

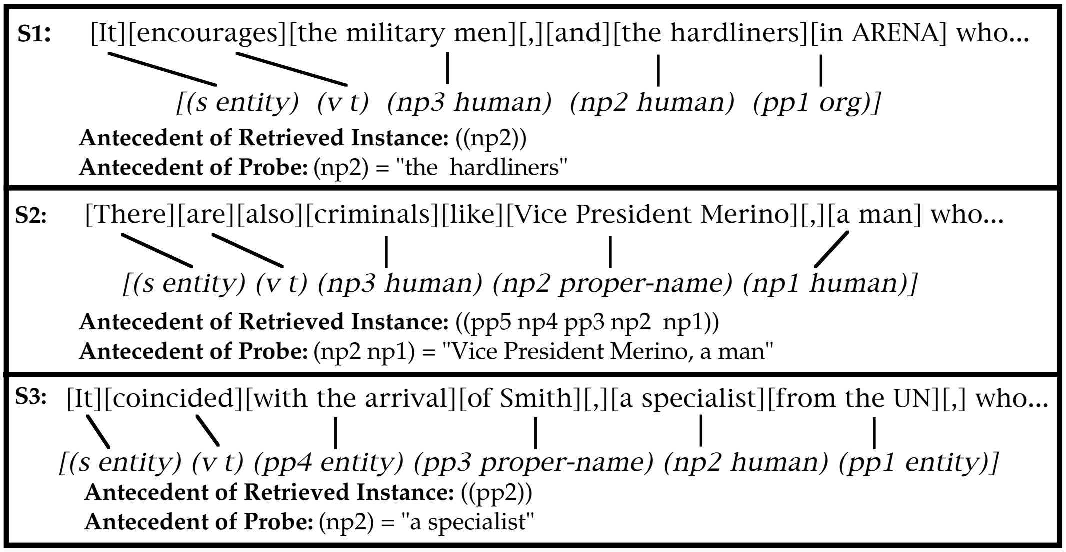

***Figure 4*. Using the Disambiguation Heuristics Hierarchy**

4 RESULTS

As described above, we used 100 texts (approximately 7% of the corpus) containing 176 instances of the relative pronoun "who" for training. Six of those instances were discarded when the UMass/MUC-3 syntactic analyzer failed to include the desired antecedent as part of its constituent representation, making it impossible for the human supervisor to specify the location of the antecedent.[7] After training, we tested the resulting disambiguation hierarchy on 71 novel instances extracted from an additional 50 texts in the corpus. Using the selection heuristics described above, the correct antecedent was found for 92% of the test instances. Of the 6 errors, 3 involved probes with antecedent combinations never seen in any of the training cases. This usually indicates that the semantic and syntactic structure of the novel clause differs significantly from those in the disambiguation hierarchy. This was, in fact, the case for 2 out of 3 of the errors. The third error involved a complex conjunction and appositive combination. In this case, the retrieved antecedent specified 3 out of 4 of the required constituents.

If we discount the errors involving unknown antecedents, our algorithm correctly classifies 94% of the novel instances (3 errors). In comparison, the original UMass/MUC-3 system that relied on hand-coded heuristics for relative pronoun disambiguation finds the correct antecedent 87% of the time (9 errors). However, a simple heuristic that chooses the most recent phrase as the antecedent succeeds 86% of the time. (For the training sets, this heuristic works only 75% of the time.) In cases where the antecedent was not the most recent phrase, UMass/MUC-3 errs 67% of the time. Our automated algorithm errs 47% of the time.

It is interesting that of the 3 errors that did not specify previously unseen antecedents, one was caused by parsing blunders. The remaining 2 errors involved relative pronoun antecedents that are difficult even for people to specify: 1) "... 9 rebels died at the hands of members of the civilian militia, who resisted the attacks" and 2) "... the government expelled a group of foreign drug traffickers who had established themselves in northern Chile". Our algorithm chose "the civilian militia" and "foreign drug traffickers" as the antecedents of "who" instead of the preferred antecedents "*members* of the civilian militia" and "*group* of foreign drug traffickers."[8]

5 CONCLUSIONS

We have described an automated approach for the acquisition of relative pronoun disambiguation heuristics that duplicates the performance of hand-coded rules. Unfortunately, extending the technique for use with unrestricted texts may be difficult. The UMass/MUC-3 parser would clearly need additional mechanisms to handle the ensuing part of speech and

---

[7] Other parsing errors occurred throughout the training set, but only those instances where the antecedent was not recognized as a constituent (and the wh-word had an antecedent) were discarded.

[8] Interestingly, in work on the automated classification of nouns, (Hindle, 1990) also noted problems with "empty" words that depend on their complements for meaning.

word sense disambiguation problems. However, recent research in these areas indicates that automated approaches for these tasks may be feasible (see, for example, (Brown, Della Pietra, Della Pietra, & Mercer, 1991) and (Hindle, 1983)). In addition, although our simple semantic feature set seems adequate for the current relative pronoun disambiguation task, it is doubtful that a single semantic feature set can be used across all domains and for all disambiguation tasks.[9]

In related work on pronoun disambiguation, Dagan and Itai (1991) successfully use statistical cooccurrence patterns to choose among the syntactically valid pronoun referents posed by the parser. Their approach is similar in that the statistical database depends on parser output. However, it differs in a variety of ways. First, human intervention is required not to specify the correct pronoun antecedent, but to check that the complete parse tree supplied by the parser for each training example is correct and to rule out potential examples that are inappropriate for their approach. More importantly, their method requires very large corpora of data.

Our technique, on the other hand, requires few training examples because each training instance is not word-based, but created from higher-level parser output.[10] Therefore, unlike other corpus-based techniques, our approach is practical for use with small to medium-sized corpora in relatively narrow domains. ((Dagan & Itai, 1991) mention the use of semantic feature-based cooccurrences as one way to make use of a smaller corpus.) In addition, because human intervention is required only to specify the antecedent during the training phase, creating disambiguation heuristics for a new domain requires little effort. Any NLP system that uses semantic features for describing nouns and has minimal syntactic parsing capabilities can generate the required training instances. The parser need only recognize noun phrases, verbs, and prepositional phrases because the disambiguation heuristics, not the parser, are responsible for recognizing the conjunctions and appositives that comprise a relative pronoun antecedent. Moreover, the success of the approach for structurally complex antecedents suggests that the technique may provide a general approach for the automated acquisition of disambiguation rules for other problems in natural language processing.

[9] In recent work on the disambiguation of structurally, but not semantically, restricted phrases, however, a set of 16 predefined semantic categories sufficed (Ravin, 1990).

[10] Although further work is needed to determine the optimal number of training examples, it is probably the case that many fewer than 170 instances were required even for the experiments described here.

6 ACKNOWLEDGMENTS

This research was supported by the Office of Naval Research, under a University Research Initiative Grant, Contract No. N00014-86-K-0764, and Contract No. N00014-92-J-1427, and NSF Presidential Young Investigators Award NSFIST-8351863 (awarded to Wendy Lehnert) and the Advanced Research Projects Agency of the Department of Defense monitored by the Air Force Office of Scientific Research under Contract No. F49620-88-C-0058.

7 REFERENCES

Brent, M. (1991). Automatic acquisition of subcategorization frames from untagged text. *Proceedings, 29th Annual Meeting of the Association for Computational Linguists*. University of California, Berkeley. Association for Computational Linguists.

Brown, P. F., Della Pietra, S. A., Della Pietra, V. J., & Mercer, R. L. (1991). Word-sense disambiguation using statistical methods. *Proceedings, 29th Annual Meeting of the Association for Computational Linguists*. University of California, Berkeley. Association for Computational Linguists.

Cardie, C., & Lehnert, W. (1991). A Cognitively Plausible Approach to Understanding Complex Syntax. *Proceedings, Eighth National Conference on Artificial Intelligence*. Anaheim, CA. AAAI Press / The MIT Press.

Correa, N. (1988). A Binding Rule for Government-Binding Parsing. *Proceedings, COLING '88*. Budapest.

Dagan, I. and Itai, A. (1991). A Statistical Filter for Resolving Pronoun References. In Y.A. Feldman and A.Bruckstein (Eds.), *Artificial Intelligence and Computer Vision* (pp. 125-135). North-Holland: Elsevier.

de Marcken, C. G. (1990). Parsing the LOB corpus. *Proceedings, 28th Annual Meeting of the Association for Computational Linguists*. University of Pittsburgh. Association for Computational Linguists.

Fisher, D. H. (1987). Knowledge Acquisition Via Incremental Conceptual Clustering. *Machine Learning*, *2*, 139-172.

Frazier, L. (1978). *On comprehending sentences: Syntactic parsing strategies*. Ph.D. Thesis. University of Connecticut.

Gluck, M. A., & Corter, J. E. (1985). Information, uncertainty, and the utility of categories. *Proceedings, Seventh Annual Conference of the Cognitive Science Society*. Lawrence Erlbaum Associates.

Hindle, D. (1983). *User manual for Fidditch* (7590-142). Naval Research Laboratory.

Hindle, D. (1990). Noun classification from predicate-argument structures. *Proceedings, 28th Annual Meeting of the Association for Computational Linguists*. University of Pittsburgh. Association for Computational Linguists.

Hindle, D., & Rooth, M. (1991). Structural ambiguity and lexical relations. *Proceedings, 29th Annual Meeting of the Association for Computational Linguists*. University of California, Berkeley. Association for Computational Linguists.

Hobbs, J. (1986). Resolving Pronoun References. In B. J. Grosz, K. Sparck Jones, & B. L. Webber (Eds.), *Readings in Natural Language Processing* (pp. 339-352). Los Altos, CA: Morgan Kaufmann Publishers, Inc.

Ingria, R., & Stallard, D. (1989). A computational mechanism for pronominal reference. *Proceedings, 27th Annual Meeting of the Association for Computational Linguistics*. Vancouver.

Kimball, J. (1973). Seven principles of surface structure parsing in natural language. *Cognition*, *2*, 15-47.

Lappin, S., & McCord, M. (1990). A syntactic filter on pronominal anaphora for slot grammar. *Proceedings, 28th Annual Meeting of the Association for Computational Linguistics*. University of Pittsburgh. Association for Computational Linguistics.

Lehnert, W. (1990). Symbolic/Subsymbolic Sentence Analysis: Exploiting the Best of Two Worlds. In J. Barnden, & J. Pollack (Eds.), *Advances in Connectionist and Neural Computation Theory*. Norwood, NJ: Ablex Publishers.

Lehnert, W., Cardie, C., Fisher, D., Riloff, E., & Williams, R. (1991).University of Massachusetts: Description of the CIRCUS System as Used for MUC-3. *Proceedings, Third Message Understanding Conference (MUC-3)*. San Diego, CA. Morgan Kaufmann Publishers.

Ravin, Y. (1990). Disambiguating and interpreting verb definitions. *Proceedings, 28th Annual Meeting of the Association for Computational Linguists*. University of Pittsburgh. Association for Computational Linguists.

Schank, R., & Riesbeck, C. (1981). *Inside Computer Understanding: Five Programs Plus Miniatures*. Hillsdale, NJ: Lawrence Erlbaum.

Sundheim, B. M. (May,1991). Overview of the Third Message Understanding Evaluation and Conference. *Proceedings,Third Message Understanding Conference (MUC-3)*. San Diego, CA. Morgan Kaufmann Publishers.

Taraban, R., & McClelland, J. L. (1988). Constituent attachment and thematic role assignment in sentence processing: influences of content-based expectations. *Journal of Memory and Language*, *27*, 597-632.

Whittemore, G., Ferrara, K., & Brunner, H. (1990). Empirical study of predictive powers of simple attachment schemes for post-modifier prepositional phrases. *Proceedings, 28th Annual Meeting of the Association for Computational Linguistics*. University of Pittsburgh. Association for Computational Linguistics.