From: AAAI-88 Proceedings. Copyright ©1988, AAAI (www.aaai.org). All rights reserved.

Acquiring Lexical Knowledge from Text: A Case Study

Paul Jacobs and Uri Zernik
Artificial Intelligence Program
GE Research and Development Center
Schenectady, NY 12301 USA

Abstract

Language acquisition addresses two important text processing issues. The immediate problem is understanding a text in spite of the existence of lexical gaps. The long term issue is that the understander must incorporate new words into its lexicon for future use. This paper describes an approach to constructing new lexical entries in a gradual process by analyzing a sequence of example texts. This approach permits the graceful tolerance of new words while enabling the automated extension of the lexicon. Each new acquired lexeme starts as a set of assumptions derived from the analysis of each word in a textual context. A variety of knowledge sources, including morphological, syntactic, semantic, and contextual knowledge, determine the assumptions. These assumptions, along with justifications and dependencies, are interpreted and refined by a learning program that ultimately updates the system's lexicon. This approach uses existing linguistic knowledge, and generalization of multiple occurrences, to create new operational lexical entries.

1 Introduction

Producing an analysis of natural language text, even in a restricted domain, requires a rich and robust lexicon. Developing such a lexicon automatically [Byrd, 1988; Boguraev and Briscoe, 1988] has emerged as one of the immediate challenges for natural language processing, because of the overwhelming difficulty of lexicon construction by hand. The automatic derivation of word meanings also improves both the relevance of the definitions to a particular domain and the internal consistency of the derived lexicon. In this paper we describe a method for acquiring new words from multiple examples in texts. We explain how the programs RINA [Zernik, 1987] and TRUMP [Jacobs, 1987] cooperate in implementing this method, and we give empirical data to motivate the entire approach.

1.1 The Process of Word Acquisition

The full acquisition of words and word meanings is a monumental task. Since partial knowledge aids a language analyzer in coping with a lexical gap and in acquiring word meanings, it makes sense to consider the gradual development of lexical entries reflecting the examples given in the texts. The following is a typical input from the domain of corporate takeovers:

```
(1)  Warnaco received another merger offer,
     valued at $36 a share.
```

If the existing lexicon covers all of the words in the input except the word **merger**, the system can still make some initial assumptions based on partial information from the text. This accomplishes the following two results:

- *Immediate result:* The sentence itself is processed and a partial meaning is produced.
- *Long-Term result:* Multiple *lexical hypotheses* are retained for processing further examples. By three competing hypotheses **merger offer** could be either: (a) an offer for a **merger** (some unknown type of transaction), (b) an offer that is **merger** (perhaps larger) than a previous known offer, or (c), an offer by a **merger** (one who performs **merging** actions).

Consider the second encounter with **merger**, as shown below:

```
(2) The merger was completed last week.
```

The interpretation of sentence (2) relies on the lexical hypotheses constructed by sentence (1). Sentence (2) provides *confirmation* for one of the lexical hypotheses (i.e., **merger** is a transaction), and rules out several others. The next two encounters are more problematic:

```
(3)  Bendix finally merged with United
     Technologies.
(4)  Midcon will be merged into a subsidiary of
     Occidental.
```

The difficulty at each such encounter is whether a current hypothesis applies, or, because of syntactic or semantic differences, the new item requires an entirely separate lexical hypothesis. In other words, since the order of the provided examples is arbitrary, the system must determine how each new example fits with prior assumptions about **merge**. Because repeated examples, as in this case, tend to be consistent with one another but to reflect somewhat different lexical structures, a system that merely treats each independently will miss important lexical information. The best approach thus seems to be to use multiple examples to derive a sufficiently general lexical hypothesis.

1.2 Issues in Language Acquisition

This approach to text processing differs from most language analyzers in the following two characteristics: (a) lexical information is taken to be incomplete, and (b) lexical knowledge is automatically upgraded as text is being processed. This involves three tasks:

Coping with Lexical Gaps: Most language processing systems either assume full lexical knowledge, or rely on user interaction to help with unknown words [Haas and Hendrix, 1983]. Our approach is more in line with that proposed by [Granger, 1977; Carbonell, 1979; Berwick, 1983; Mooney, 1987]: to use linguistic and conceptual context to "guess" the general category of a new word. For example, the word combination **another merger offer** is problematic due to the absence of either **merger** or **merger offer** in the lexicon, but it can be processed using knowledge of other similar cases such as **buyout offer** and **settlement offer**, and the more general **specifier noun1 noun2** compound.

Forming a Hypothesis: In order to complete a parse in the presence of an unknown word, and to posit a meaning for the word, the system must take into account a variety of linguistic and conceptual clues. For example, the **-er** ending suggests either a comparative adjective or nominal suffix, each of which has conceptual content. At this point, having received a single example, the system must entertain many competing assumptions that are all very skeletal. In this case, it is not difficult to tolerate the unknown word, but there is not enough information to create a new lexical entry from the single example.

Refining the Hypothesis: Further examples are required to prune the space of hypotheses above and to add enough detail to handle related cases, such as those in examples (2-4). A generalization process results in the creation of a new lexical entry from multiple examples.

1.3 Alternative Methods in Lexical Acquisition

The previous section has identified three different aspects of lexical acquisition: (1) determining word or phrase meaning from context, (2) determining a composite meaning by combining constituents, and (3) using phrasal structure to isolate the meaning of a single word in the phrase.

Programs such as FOULUP [Granger, 1977] and CHILD [Selfridge, 1980] perform the first task, using context to determine meaning. The meaning of the word **waitress**, for example, is acquired by identifying that particular role in the restaurant script. Learning involves the association of the new word with the salient element in the context. Three assumptions are implicit in this approach: (a) the context is fully known and well represented, (b) morphological or syntactic information is unnecessary or secondary, and (c) the given example is perfect and so learning is a one-shot process. This approach is simplistic, as it ignores the language as a knowledge source.

The second approach takes linguistic clues into account. The morphology of the word **waitress**, for example, includes semantic information: it indicates **waiting**, and it identifies the gender of the person. Moreover, a prediction can be made about the possible existence of a male counterpart called **waiter**. Programs such as RINA [Zernik, 1987] and GENESIS [Mooney, 1987] have deduced the meaning of whole phrases from their parts. However, not all English phrases are *productive*: the meanings of non-productive phrases such as **to take somebody on** or **the big apple** cannot be derived from the meanings of their constituents. The meaning of a single word, such as **merger**, also cannot be determined easily from the meaning of the root (**merge**) and morphological knowledge. Meanings of such phrases must depend on context. Thus it is important to consider morphology along with phrasal and conceptual context in determining an applicable meaning.

The third issue mentioned above, isolating the meaning of an individual word, reflects an often-overlooked aspect of the acquisition process. While the first two aspects assume the computation of a whole meaning from the meanings of constituents, in many cases it is the constituent that must be acquired. As shown in sentence (1), the meaning of the entire noun phrase **another merger offer** is known from the context (a similar offer has been previously introduced), but the meaning of the word **merger** is yet unknown. Thus the meaning of the part must be extracted from the meaning of the whole phrase. This problem pervades texts where new modifiers are frequent, and the contribution of the modifier itself must be extracted from the entire concept.

Our program combines these three approaches. In particular, this report shows how the meaning of a modifier is extracted from the meaning of an entire phrase (using method 3), which is derived from prior discourse (method 1) and by the combination of linguistic clues (method 2).

1.4 Empirical Observations

The prototype system SCISOR [Rau, 1987] processes and retrieves information from news stories about corporate takeovers. Here we describe its text processing (TRUMP) and acquisition (RINA) components. TRUMP [Jacobs, 1986; Jacobs, 1987] analyzes input text in the presence of partial information. RINA, a language learning program [Zernik, 1987; Zernik and Dyer, 1988], performs validation and generalization. RINA previously learned by interaction with a user. In this research it is being adapted to learn from textual examples alone.

SCISOR is designed to process information in a single domain. Its base lexicon includes 10,000 roots that cover most of the more frequent and more general vocabulary. The "new" words (those that cannot be successfully analyzed using the roots and basic morphology) tend, therefore, to have a specialized meaning that can be related to structures in the base lexicon.

The appendix shows some simple examples drawn from a corpus of about 40,000 words coming over the Dow Jones News Service on January 25, 1988. On this typical day, 33 new words occurred at least 7 times within the corpus. These words can be roughly divided into three groups: (A) those whose syntactic and semantic behavior could be determined successfully using a more complete morphological analysis (e.g., the meaning of **unchanged** can be produced from its root and its prefix, since **change** already exists in the lexicon), (B) words that can be easily categorized syntactically but not semantically (e.g., **debentures**, based on its morphology, is identified as a noun), and (C) those for which relatively little information is available without looking at the surrounding context (e.g., **crude** cannot be identified by either a syntactic or a semantic category).

The third group is generally the hardest, but the data reveal that the commonly occurring words in this group are used in fairly limited contexts. This phenomenon gives power to an approach that relies on context to generalize a hypothesis about a new word.

The examples of words in the third group, such as **crude**, **dividend**, and **pence**, show that the common phrases (e.g., **crude oil**) in which these new words occur often reveal the meaning of the individual word. **Merger** is among the more difficult words in the third group, because the morphology is unhelpful and in fact misleading in determining the syntax and meaning of the unknown word. In this case, the syntactic and conceptual context allow SCISOR to produce a hypothesis for the new lexical entry.

The rest of this paper describes the interaction of the text-processing components. TRUMP receives the input text, and produces a *hypothesis chart*—a structure denoting the input text, embedded in which are *assumptions*—possible characteristics of each unknown word. RINA uses this hypothesis chart to generate *lexical hypotheses* consistent with the local context. These hypotheses are retained in the lexicon, and are refined, or withdrawn, during multiple occurrences of a word or word root.

2 Forming Assumptions

The input to TRUMP is natural language text; the output from the system is a data structure, or chart, representing the space of possible syntactic and semantic interpretations. When the system encounters unknown words, TRUMP uses assumptions about the unknown words to construct the chart. The following knowledge sources contribute to these assumptions:

1. *Morphology:* **Merger** has several morphological breakdowns, each conveying some conceptual content. **Merger** can be either a simple root (like **ledger**); a comparative adjective with root **Merg-** or **Merge-** (like **larger**); or a noun (like **shopper**, **buyer** and **seller**), where **merger** describes the actor of some hypothetical action—the yet unknown **merge/merg** action in this case.
2. *Syntax:* The word may be an adjective or noun serving as a modifier, and may be part of the compound nominal **merger offer**.
3. *Semantics/Conceptual:* **Merger offer** may describe a type of offer or something being offered. It may also describe an offer given by someone who **merges**.
4. *Semantics/Lexical:* **Received ... merger offer** suggests that a **merger offer** was the object of an abstract transfer event in which **Warnaco** was the recipient.
5. *Context:* The specifier **another** presupposes the existence of a **merger offer** in the context. Thus, either there was an offer before and this offer is **merger** than that one, there was another **merger offer** before, or there was an offer and this is one of a group of offers that are **merger** than the first.

In the following subsections we explain the contents of single assumptions and the structure of the entire hypothesis chart.

2.1 Generating Assumptions

For a lexical gap, such as an unknown word, TRUMP produces an ordered list of candidate interpretations, including morphological and syntactic characteristics. These are then used in completing the syntactic/semantic chart in much the same manner as in the case of known words.

For **merger**, the result of the initial analysis is a set of candidates, as shown in figure 1. Braces in the figure group together sets of mutually exclusive assumptions. The main assumptions are labeled **H1**, **H2**, or **H3** in the diagram. Thus, **merger** is either a comparative adjective meaning more *merge* (**H2**), a noun (**H1**), or a single adjective (**H3**) in this linguistic context. If the word derives from a root plus an **-er** ending, the root can be either **merge-** or **merg-**.

2.2 Producing a Hypothesis Chart

The previous section described the analysis of a single unknown word in a linguistic context. The next step is to build from this analysis an interpretation of an entire sentence. TRUMP constructs a semantic representation of its input by mapping linguistic structures into conceptual structures and combining the new structures. The hypothesis chart includes the possible conceptual interpretations along with the assumptions on which they depend. For example, the hypothesis chart for part of sentence (1) is shown in figure 2.

The linguistic assumptions **H1**–**H3** in figure 2 are identical to those in figure 1, thus the hypothesis chart includes the word analysis. Conceptual assumptions, such as that *merge* is a quality and that **merger offer** is thus more *merge*, are dependent on corresponding linguistic assumptions (labeled **H3** in this case). These assumptions in turn depend on the morphological analysis shown in the previous diagram. When a conceptual assumption proves false, the corresponding linguistic assumptions can be corrected because they are explicitly related within the chart. Similarly, support for a conceptual assumption also supports a corresponding linguistic assumption. The main advantage of this approach is that it allows multiple sources of linguistic knowledge to contribute to the language acquisition process. This process is described in the next section.

3 Forming Lexical Hypotheses

From the assumptions introduced by TRUMP, RINA generates full-fledged lexical hypotheses associating structural forms and their meanings. The following sections describe the structures received by RINA in this example and one process by which an operational lexical entry is generated.

3.1 The Input

Initially RINA receives a pair: a *chart* representing the text; and a *context*, presenting the objects referred to in prior discourse.

HYPOTHESIS CHART

```
subject (head-noun warnaco
                concept -> company.34)

verb    (root receive
```

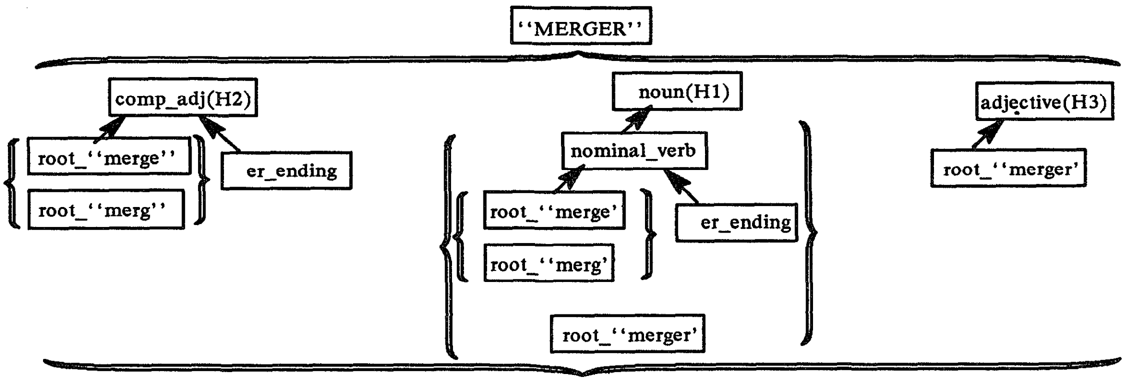

Figure 1: Word analysis chart of **merger**

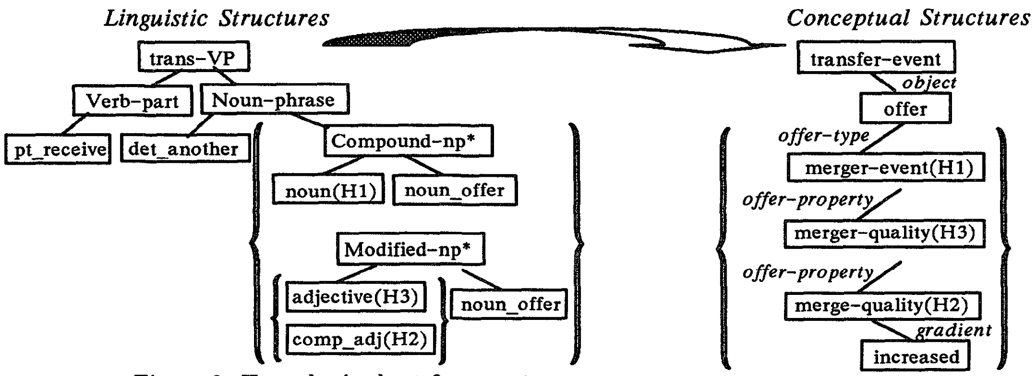

Figure 2: Hypothesis chart for **received another merger offer**

```
          concept -> comm-transfer)

object (specifier another
        modifier merger head-noun offer
              concept -> XX)
```

PRIOR CONTEXT

```
company.34:
  (company :name warnaco)
company.46:
  (company :name wac)
offer.123:
  (offer :offerer company.46
         :offeree company.34
         :offered
            (company-acquisition
               :acquirer company.46
               :acquired company.34))
```

The top structure represents one particular path in the given chart; the three concepts underneath represent the part of the relevant context prior to reading sentence (1).

3.2 Stages in Hypothesis Formation

Using these two structures, RINA creates new lexical entries in four steps: (a) scoping the unknown element in the given chart, (b) converting that element from an instance to a general template, (c) forming a lexical hypothesis based on a set of assumptions, and (d), refining the hypothesis by further examples.

Scoping the Unknown: Within the given input, RINA must determine the scope of the unknown lexical entry. RINA starts learning by focusing on an entire *reference-referent* association:

Template1:

```
reference: another merger offer
referent : offer.123--WAC offer to acquire
           Warnaco
```

RINA associates the incomplete reference **another merger offer** with *offer.123*, in spite of the unknown word **merger**, by using three clues: (a) the word **offer**, which yields a basic description of the event; (b) the determiner **another**, which suggests a prior offer in the context; and (c) the existence of *offer.123* in the context.

Initial Generalization: Although the association above provides a pattern-meaning pair, it still presents a specific instance, and it must be converted into the following operational lexical entry:

Template2:

```
pattern: merger offer
concept: X offer Y to acquire company Y
```

The conversion from an instance to a template (called *variabilization*) is accomplished in two steps: (a) **another** is identified as a general specifier which should be removed (**merger offer** must not always be preceded by **another**); (b) the conceptual template is "uniquefied": names that are associated with this particular context are uniformly replaced by variables.

Assumption-Based Generalization: Template2 above does not capture the full behavior of **merger**. For robust coverage, RINA must exploit the assumptions given by TRUMP in constructing further hypotheses. We describe here one particular path, in which three given assumptions are used:

(a) **Merger** is a noun, root **merger**;

(b) In the compound **merger offer**, the noun **merger** is a modifier, modifying the noun **offer** .

(c) **Merger offer** is an instance of a general lexical entry **transaction offer**, which denotes *an offer to perform a transaction.*

Through these assumptions, RINA interprets **Template1** above as a productive composition of its constituents:

Template3:

```
pattern: <:modifier merger :head-noun offer>
concept: offer to perform a merger-transaction
```

RINA uses structure matching with the productive pattern of **Template3**, to break down **Template1** into its constituents. Accordingly, **merger** is associated with *company acquisition*, leading to the following lexical hypothesis:

Template4:

```
pattern: merger
concept: company-acquisition
```

The templates above play two important roles. **Template2** helps to construct a meaning for sentence (1). **Template4** represents a long-term hypothesis to be employed in future examples.

Hypothesis Refinement: RINA's hypotheses are further refined by each of the following examples:

```
(2)  The merger was completed last week.
(3)  Bendix finally merged with United
     Technologies.
(4)  Midcon will be merged into a subsidiary of
     Occidental.
```

Sentence (2) provides a *confirmation* of the hypothesis above (and rules out other spurious hypotheses that are not described here); Sentence (3) initiates a new assumption. The morphological relation between **merger** the noun, and **merge/merg** the verb is used for transfering semantic properties of the already known noun to the yet unknown verb. Accordingly, **to merge** means *perform company acquisition.* This hypothesis misses the general meaning of **to merge**— by ignoring the general nature of merger as an act of integration and by assuming incorrectly that a merger is an asymmetric act—but this notion of merger highlights the way words are customized as special terms in particular domains.

4 Conclusions

TRUMP and RINA, the text processing components of SCISOR, are written in CommonLisp and run on the SUN workstation. The program implemented so far can handle only few examples. Its lexicon, in particular the morphological database, must be extended in order to exploit a wider set of linguistic clues; the learning algorithm must be generalized to handle a variety of input sequences.

Theoretical Limitations: This learning method, designed to operate without explicit user teaching, is sensitive to the input text in two ways: (a) the order of examples, and (b) the linguistic clues presented in each example, have a strong impact on the resulting concepts.

Here, the word **merger** was acquired without prior knowledge of the general verb **to merge**. As a result, the acquired meaning is undergeneralized, and skewed toward the corporate takeover domain. A different sequence yields different results. When the general verb **to merge** exists a priori in the lexicon, the term **merger** is interpreted as a specialization of that verb. Thus, learning by this method is order dependent.

Learning of the **merger** concept also exploited the special role of the determiner **another**. Accordingly, the new concept was related to a concept already existing in the context. Since this program does not rely on explicit tutoring (i.e., **a company merger may be a company acquisition**), it must systematically utilize clues such as modifiers, determiners, and connectives which are pervasive in the provided text. Thus, the acquired concepts reflect the clues found in each particular example.

Summary: It is unrealistic to assume that a program's linguistic database can be complete. Therefore a robust program must be able to handle cases where linguistic knowledge is missing. Moreover, it must acquire that knowledge from the texts themselves. While there is no general method for acquiring general word meanings from texts, the approach described here helps both in coping with lexical gaps and in forming lexical hypotheses that are suitable for texts within a domain. This approach combines the assumption-based language analysis strategy of TRUMP with the hypothesis refinement of RINA. Lexical information is acquired in a gradual process from multiple examples in the text.

A Examples from a Corpus

The following table includes the 33 new words on January 25, grouped roughly according to level of morphological analysis. Words in Group C are difficult since in isolation, no morphological clues are found as to their syntactic or semantic categorization.

A	B	C
unchanged	debentures	crude
takeover	savings	dividend
holders	headlines	pence
restructuring	implications	roundup
stockholders	sustained	definitive
decliners	volatility	downgrade
repurchase	speculation	arbitrage
convertible	rejected	imbalance
buyout	sweetened	subordinated
institutional	refractories	uncertainty

For the three most common words in column C, the following are the surrounding contexts extracted from the original texts:

```
OMPANY EXPECTS THAT CRUDE OIL PRICES MAY REMAIN S
-25-88:"?;  IRANIAN CRUDE OUTPUT AVERAGED 2 MILLI
MANAMA -DJ- IRANIAN CRUDE-OIL OUTPUT DURING THE F
DICATE THAT IRANIAN CRUDE EXPORTS ARE CURRENTLY A
 MILLION BARRELS OF CRUDE STORAGE IN VESSELS IN T
7'S FOURTH-QUARTER, CRUDE PRICES SOFTENED AND CRU
PRICES SOFTENED AND CRUDE SUPPLY COSTS TO FOREIGN
OMBINATION OF LOWER CRUDE SUPPLY COSTS AND RELATI
 STRENGTH OF HIGHER CRUDE OIL PRICES.      FOURTH
PACT OF HIGHER 1987 CRUDE OIL REALIZATIONS WAS MO
 STRENGTH OF HIGHER CRUDE PRICES AND COST SAVINGS
 EST 01-25-88:"?;   CRUDE OIL FUTURES LOWER; MARC
:'    NEW YORK -DJ- CRUDE OIL FUTURES PRICES ARE
 EXCHANGE.    MARCH CRUDE IS QUOTED AT $16.91 A B
6.66 AND $16.84.    CRUDE FUTURES PRICES, WHICH O
ALLING JUST SHORT.  CRUDE PRICES THEN SANK BACK T
 SPOT-MARKET PRICED CRUDE DEAL WITH A JAPANESE CO
 A RESULT OF HIGHER CRUDE PRICES AND LOWER EXPLOR
 COMPLETELY RECOVER CRUDE OIL PRICE INCREASES IN
INGS REFLECT HIGHER CRUDE OIL PRICES AND PRODUCTI
ARNINGS WERE HIGHER CRUDE OIL PRICES AND INCREASE
:'    NEW YORK -DJ- CRUDE OIL FUTURES SETTLED HIG
ANTILE EXCHANGE.    CRUDE OIL FOR MARCH DELIVERY
8 AND $17.12. APRIL CRUDE WAS AT $16.96, ALSO UP
MINUTES OF TRADING, CRUDE PRICES, WHICH HAD FAILE
7 A BARREL IN MARCH CRUDE FADED WHEN PRODUCT
N IT PAID A SPECIAL DIVIDEND OF $33 A SHARE TO TH
 COMPANY CALLED THE DIVIDEND AN INITIAL QUARTERLY
N INITIAL QUARTERLY DIVIDEND.     -0-;  9 54 AM E
 SUNSTATES DECL STK DIVIDEND        JACKSONVILLE,
ARED A 100 PC STOCK DIVIDEND ON ITS COMMON STOCK
 FEB. 1 SEMI-ANNUAL DIVIDEND ON ITS $3.75 CUMULAT
ING ON TAKEOVER AND DIVIDEND PLAYS.    THE DOW JO
 STOCKS WHICH GO EX-DIVIDEND TUESDAY - SOUTHERN C
BE IMPLEMENTED BY A DIVIDEND OF ONE PREFERRED STO
 SOME TRADERS SAID. DIVIDEND-CAPTURE STRATEGIES A
EARLIER GAINS, WITH DIVIDEND-RELATED TRADES AND T
ING STOCKS GOING EX-DIVIDEND TUESDAY ACCOUNTS FOR
K CARRIES AN 8.8 PC DIVIDEND YIELD AND GOES EX-DI
D YIELD AND GOES EX-DIVIDEND TUESDAY.    THE NEXT
RE ALSO INVOLVED IN DIVIDEND PLAYS - PINNACLE WES
:'    NEW YORK -DJ- DIVIDEND-RELATED TRADES IN A
A REGULAR QUARTERLY DIVIDEND OF 11 CENTS, PAYABLE
 STOCK TRADERS SAID DIVIDEND CAPTURE STRATEGIES A
LCARES   5 PC STOCK DIVIDEND                 PAY M
 AS A 2-FOR-1 STOCK DIVIDEND FOR EACH SHARE OF CL
LCARES   5 PC STOCK DIVIDEND                 PAY M
AID A SIMILAR STOCK DIVIDEND IN APRIL. -0-;  4 08
FRIDAY DUE TO BRISK DIVIDEND-RELATED ACTIVITY IN
E THE IMPACT OF THE DIVIDEND PLAYS ON VOLUME, HE
 DECLARED QUARTERLY DIVIDEND OF 7C WITH THE SAME
K CARRIES AN 8.8 PC DIVIDEND YIELD AND GOES EX-DI
D YIELD AND GOES EX-DIVIDEND TUESDAY.     THE NEX
SO WERE INVOLVED IN DIVIDEND PLAYS - PINNACLE WES
ETENED OFFER OF 500 PENCE A SHARE FOR ARCO'S STAK
 ALREADY OWN TO 500 PENCE A SHARE, OR 1.77 BILLIO
ON POUNDS, FROM 450 PENCE, OR 1.59 BILLION POUNDS
OIL'S ASSETS AT 699 PENCE A SHARE.    BRITOIL ALS
L SHARES WERE UP 19 PENCE AT 478, WHILE BP SHARES
IL, WHICH WAS UP 20 PENCE AT 479 AFTER BRITISH PE
 THE COMPANY TO 500 PENCE A SHARE FROM 450 PENCE.
CE A SHARE FROM 450 PENCE.  BRITOIL REJECTED THE
AS UNCHANGED AT 259 PENCE.    ROLLS-ROYCE PLC SAI
 SHARES WERE DOWN 7 PENCE AT 93.    TI GROUP COMP
TI SHARES WERE UP 3 PENCE AT 335.    AMONG OTHER
S, SHELL WAS DOWN 8 PENCE AT 10.20 POUNDS, ICI DO
 POUNDS, ICI DOWN 9 PENCE AT 10.74 POUNDS, GLAXO
UNDS, GLAXO DOWN 17 PENCE AT 10.14 POUNDS, JAGUAR
RCHASES WERE AT 350 PENCE PER SHARE.    ''HOWEVER
TENDER OFFER AT 450 PENCE PER SHARE FOR ALL REMAI
TENDER OFFER TO 500 PENCE PER SHARE.'' THE RATING
 FOR A PRICE OF 475 PENCE PER SHARE, AND THE MINE
```

References

[Berwick, 1983] Robert Berwick. Learning word meanings from examples. In *Proceedings of the Eighth International Joint Conference on Artificial Intelligence*, Karlsruhe, Germany, 1983.

[Boguraev and Briscoe, 1988] B. Boguraev and T. Briscoe. Large lexicons for natural language processing. *The Journal of Computational Linguistics, The Special Issue on the Lexicon*, 1988.

[Byrd, 1988] R. et al Byrd. Tools and methods for computational lexicology. *The Journal of Computational Linguistics, The Special Issue on the Lexicon*, 1988.

[Carbonell, 1979] J. Carbonell. Towards a self-extending parser. In *Proceedings of the 17th Meeting of the Association for Computational Linguistics*, pages 3–7, 1979.

[Granger, 1977] Richard Granger. Foulup: a program that figures out meanings of words from context. In *Proceedings of the Fifth International Joint Conference on Artificial Intelligence*, 1977.

[Haas and Hendrix, 1983] N. Haas and G. Hendrix. Learning by being told: acquiring knowledge for information management. In R. Michalski, J. Carbonell, and T. Mitchell, editors, *Machine Learning: An Artificial Intelligence Approach*, Tioga Press, Palo Alto, California, 1983.

[Jacobs, 1986] Paul S. Jacobs. Language analysis in not-so-limited domains. In *Proceedings of the Fall Joint Computer Conference*, Dallas, Texas, 1986.

[Jacobs, 1987] Paul S. Jacobs. A knowledge framework for natural language analysis. In *Proceedings of the Tenth International Joint Conference on Artificial Intelligence*, Milan, Italy, 1987.

[Mooney, 1987] Raymond Mooney. Integrating learning of words and their underlying concepts. In *Proceedings of the Ninth Annual Conference of the Cognitive Science Society*, Seattle, WA, Lawrence Erlbaum, Associates, Hillsdale, NJ, July 1987.

[Rau, 1987] Lisa F. Rau. Knowledge organization and access in a conceptual information system. *Information Processing and Management, Special Issue on Artificial Intelligence for Information Retrieval*, 23(4):269–283, 1987.

[Selfridge, 1980] Mallory G. R. Selfridge. *A Process Model of Language Acquisition*. PhD thesis, Yale University Department of Computer Science, 1980.

[Zernik, 1987] Uri Zernik. Language acquisition: learning phrases in a hierarchy. In *Proceedings of the Tenth International Joint Conference on Artificial Intelligence*, Milan, Italy, 1987.

[Zernik and Dyer, 1988] U. Zernik and M. Dyer. The self-extending phrasal lexicon. *The Journal of Computational Linguistics, The Special Issue on the Lexicon*, 1988.