Lexicon Acquisition: Learning from Corpus by Capitalizing on Lexical Categories

Uri Zernik
Artificial Intelligence Program
General Electric - Research and Development Center
Schenectady, NY 12301

Abstract

Text examples must be exploited in the acquisition of lexical structures. However, neither syntactic nor semantic features are provided by the text itself, and so acquisition must be aided by additional resources. We investigate the application of an existing resource, a set of lexical categories, as a prediction method. We present an algorithm that applies (a) top-down prediction based on lexical categories; (b) bottom-up validation by scanning text examples. Finally, we discuss the issue of semantic bootstrapping and identify its theoretical and practical limitations.

1 Introduction

Existing programs frequently stumble when encountering a new word. Such lexical gaps diminish the utility of natural language technology. The problem is aggravated by the existence of entire unknown phrases composed of single well-known words:

(1) John made a table from raw wood.
(2) He made a widow happy.
(3) He made a happy widow.
(4) He made the widow a table.
(5) He made her leave early.

Each *make* phrase interacts with its arguments in its own idiosyncratic way. Example (1) presents the simple usage of make: make means generate. Examples *(2)* and (3), both taken from *Love in the Time of Cholera* [Marquez, 1986], are more intriguing since similar words combine in entirely different ways: in (2) she (the widow) becomes happy; in (3) he becomes a happy widow. Example (4) introduces the *beneficiary* interaction: he made it *for* her. And finally example (5) brings in the complement-taking form: he forced her to act. These examples illustrate how subtle differences in argument structure might impact the entire meaning. A lexicon therefore must account for all the variations such a verb can possibly assume. Idiomatic phrases are not confined to fine literature. The sample sentences next page are taken from the Dow-Jones newswire (July 7, 1988).

A brief observation reveals the diversity of phrases used in this technical domain. *Make a statement, make a plan,* and *make a decision* fall into one category. *Make final net $10,000* falls into a second category. *Make it difficult, make it attractive, make it available* fall into yet another category. This small collection of sentences illustrates (a) how extensive a lexicon should be to facilitate effective text processing, and (b), the wealth of raw information provided in the text for lexical acquisition purposes.

A program processing such text cannot be provided with all these categories at the outset. Therefore, lexical knowledge must be acquired on demand: once a lexical unknown is encountered while processing an individual sentence, the program should extract the new entry from available resources. The newswire text itself, and other on-line corpora are readily available and should be used.

In this paper we describe a learning algorithm which exploits on-line text and existing lexical categories. The algorithm applies top-down prediction based on lexical categories. Since the correlation between syntax and semantics provided by a lexical category is limited, the predictions must be validated by bottom-up scanning of text examples. We describe the algorithm and evaluate its merits.

2 Task Description

Learning programs in general are designed to enhance the performance of associated performing programs. Lexical acquisition in particular supports language processing.

2.1 Text Processing: Encountering a Lexical Gap

The existence of lexical gaps is manifested as inaccuracy in text processing. Consider the operation of the program TRUMP [Jacobs and Rau, 1988] as it processes the following sentence (taken from the Dow-Jones examples):

(6) Further modifications could make the project economically attractive.

TISFIED THE ASSETS EXPECTS TO MAKE A FURTHER STATEMENT THIS WEEK.
THE AGENCY SAID THAT WOULD MAKE IT THE MOST EXPENSIVE TOXICS CLEANUP JOB
IFICATIONS WOULD BE NEEDED TO MAKE THE PROJECT ECONOMICALLY ATTRACTIVE AND
NEW CAPITAL REQUIREMENTS WILL MAKE IT DIFFICULT TO IMPLEMENT CONSTRUCTIVE N
NY ALSO SAID IT WAS UNABLE TO MAKE A SCHEDULED DEBT PAYMENT TO ITS SAVINGS
M EST 02-01-88:"?; BRAZIL TO MAKE $350 MILLION INTEREST PAYMENT
THE GOVERNMENT OF BRAZIL WILL MAKE AN ADDITIONAL PAYMENT TOMORROW, OF ABOUT
THE GOVERNMENT HAD CHOSEN TO MAKE THE JAN. 15 PAYMENT, IN LIEU OF OTHER PA
H WILL BECOME OF EQUIPMENT IT MAKES AVAILABLE FOR CUSTOMERS TO LEASE. ON
A TAX CREDIT OF $13.5 MILLION MADE FINAL NET AVERAGE SHARES 87 MILLION V
AND A $3.9 MILLION TAX CREDIT MADE FINAL NET $117.6 MILLION SALES $2,927
D DEBENTURES DUE 1995 WILL BE MADE IN ADVANCE OF THAT DATE, CLARKE SAID.
FROM EARLY RETIREMENT OF DEBT MADE FINAL NET $1,209,000,000 OR $4.70. I
UED OPERATIONS OF $39,000,000 MADE FINAL NET $1,064,000,000 OR $3.83. SA
THE TENDER OFFER IS BEING MADE UNDER AN AGREEMENT AND AND WITHDRAWAL RI
L COURT HERE, WHERE IT CAN BE MADE FINAL AFTER A 45-DAY COMMENT PERIOD.
LIVERIES OF A MECHANICAL FUZE MADE BY GENERAL DEFENSE'S HAMILTON TECHNOLOGY
A TAX BENEFIT OF $3,545,000, MADE FINAL NET $5,092,000. SALES $136,27
HE OFFERING IS EXPECTED TO BE MADE IN LATE FEBRUARY. OF THE SHARES TO BE
DINARY TAX CREDIT OF $554,000 MADE FINAL NET $2,615,000 OR 62C. IN THE
RAORDINARY CREDIT OF $304,000 MADE FINAL NET $650,000 OR 36C A SHR. SALE
THE OFFER OF BOARD SEATS WAS MADE TO OLYMPIA 4 YORK OVER THE WEEKEND, WHIC
OF THAT LOSS THROUGH PROFITS MADE BY TRADING THE BILLS. ALTHOUGH THERE
IN BROADCASTING CORP. SAID IT MADE FILINGS WITH THE FEDERAL COMMUNICATIONS
MERICAN CELLULAR, AND SAID IT MADE TODAY'S FILINGS IN ORDER TO FACILITATE I
DCASTING CORP., WHICH SAID IT MADE FILINGS WITH THE FCC, THE DEPARTMENT OF
HE $9-A-SHARE OFFER WAS FIRST MADE IN LATE NOVEMBER AND HAS SINCE BEEN EXTE
ONOMIC PROGRESS IS ALSO BEING MADE BY MEXICO AND VENEZUELA, WHO BENEFITED F
4P SAID THE "DOWNGRADES WERE MADE WHERE THE COMBINATION OF LARGE LESSER DE
AIN RECOMMENDATIONS HAVE BEEN MADE TO AUGMENT OR IMPROVE SEVERAL ASPECTS OF
UIREMENTS AND THAT IT HAD NOT MADE A PAYMENT TO A GROUP OF SECURED BANK LEN
0 FROM EXTINGUISHMENT OF DEBT MADE FINAL NET $72,020,000 OR $2.17. RJEVE
OM CONTINUING OPERATIONS WERE MADE TO THE U.S. GOVERNMENT: 81 PC ON DEFENS
OM CONTINUING OPERATIONS WERE MADE TO THE U.S. GOVERNMENT: 81 PC ON DEFENSE
SAID THESE PURCHASES WILL BE MADE FROM TIME TO TIME AND MAY AGGREGATE UP T
MOUNTAIN, ALTHOUGH IT HAS NOT MADE ANY DEFINITE PLANS TO DO SO. HORN 4 H
KING'S TABLE INC. THAT IT HAS MADE NO DECISION AS TO WHETHER IT WILL ENHANC
BANKERS CLUB IN LONDON, POEHL MADE THE STRONGEST APPEAL FOR A DOLLAR STABIL

Figure 1: Sample MAKE Sentences on the Dow-Jones Newswire

TRUMP's lexicon at this point includes only the basic entry for make (Phrase 1 below).

structure: X.agent make Y:object

meaning: X created object Y

MAKE - phrasel

This entry is tailored for parsing sentences such as (7) and (8) below.

(7) Mary made a table from raw wood.
(8) John made a great meal.

In the absence of the appropriate phrase (Phrase2 below)

structure: X:agent make Y:object Z:attribute

meaning: X caused Z to be attributed to Y.

MAKE - phrase2

TRUMP applies Phrasel and consequently it obtains misleading results: modifications create a new project. This interpretation entails three incorrect facts:

1. a new project object is incorrectly instantiated in the program;
2. the main act is the creation of that project;
3. *modifications* are taken as agent and not as cause of that act.

Since TRUMP's output is used in propagating further inference, this interpretation might lead to inappropriate conclusions.

2.2 The Learning Algorithm

The missing phrase could be provided by a learning algorithm. The specifications of that algorithm are determined by availability of resources.

- *The input:* a sentence including an unknown lexical entry.
- *The given:*
  1. a corpus of syntactically analyzed sentences;
  2. a set of existing lexical categories.
- *The output:* classification of the unknown phrase within an existing category.

We have allowed the use of corpus (i.e., a large collection of sentences accessible on line) and lexical categories. We have explicitly precluded the use of other resources:

- On line dictionaries
- Context: the context in which each sentence appears is not given. It is unrealistic to assume that this information can be provided systematically.
- Semantics: the semantic interpretation each sentence in the corpus is not given. It is unrealistic to assume that a text processor is capable of converting sentences into semantic templates for an arbitrarily large corpus. Semantic text processors are confined by their nature to limited domains.

We did assume that sentences are syntactically analyzed as to lexical arguments. This assumption is not-trivial when considering the corpus size. Currently, our input sentences were not processed automatically. Yet in the long run, the validity of this approach will depend on this assumption - that syntactic analysis will be available for large corpus. We discuss this issue in the concluding section.

2.3 The Acquired Lexical Entry

Four elements constitute a phrase in the lexicon. The learning program must acquire all four.

Syntactic breakdown: In the given example *(Further modifications could make it economically attractive to clients),* the verb phrase includes 5 different clauses. The lexical analysis of this phrase is given below:

subject:	X:NP	;further modifications
verb:	make	;could make
objectl:	Y:NP	;it
object2:	Z:AP	:economically attractive
adjunct:	W:PP	;to clients

Within this group of clauses it is important to distinguish between a mandatory *complement* (subject, verb, objects) vs. an auxiliary *adjunt.* Here, for example *economically attractive* is a complement while *to clients* is an adjunct. Adjuncts should be identified and factored out since they do not belong in the lexical definition. [1]

Semantics: What is the basic act, and what are the thematic relations among the arguments? The semantic template is expressed in terms of the components above:

act state-change
cause X
object Y
new-attribute Z
causing-act MAKE

Variants: What are all the different configurations in which the syntactic arguments can be organized? Are there any other possible lexical variants for this phrase? Further legitimate variants of Phrase2 could be:

1. make (NP, NP, AP); (i.e., This made Mary happy)
2. make (NP, AP, NP); (i.e., Let's make public this new data!)
3. make (NP, NP, NP); (i.e., This made her one happy woman)
4. make (NP, AP, NP); (i.e., It was made available to us)

[1] The lexical breakdown above is abbreviated later as follows: (NP, NP, AP). Notice that the adjunct is not included.

Lexical and grammatical variants must be distinguished: lexical variants (1,2,3) must be recorded in the lexicon since they present properties of the phrase itself. On the other hand grammatical variants such as passive voice (4) can be derived by general grammar rules and need not be explicitly recorded. An important property of a phrase, called the phrase stamp, is the set of all lexical variants of that phrase: {(NP, NP, AP) (NP, AP, NP) (NP, NP, NP)}.

Categories: By observing sets of verbs such as *paint, cut,* and *dig* we realize that syntactic and semantic features are common to the entire set (i.e., *cut it short, dig it deep, paint it green),* which can therefore be clustered into a category. Lexical categories can readily contribute properties to yet unknown verbs. Therefore, a phrase should be indexed, if possible, within an existing category. For example, it is necessary to determine whether "make it attractive" belongs in any of the following categories:

Dative: (e.g., John gave Mary a book)
verb(NF, NP1, NP2) or verb(NP, NP2, toNPl)

Selection: (e.g., They elected Salinas president)
verb(NP, NP1, NP2) or verb(NP, NP1, to-be-NP2)

State-change: (e.g., He painted the window brown)
verb(NP, NP, AP) or verb(NP, AP, NP)

Indexing to a category is crucial since it supports prediction of rough meaning and thematic relations. Moreover, a category determines a stamp, the set of lexical variants shared by its members.

3 Theoretical Issues

Three theoretical issues must be resolved to enable learning from text.

3.1 Identifying a Lexical Gap

Unless a discrepancy is detected a lexical gap is not identified and learning could not be triggered. In section 2 we showed how a parser applies an inappropriate phrase in parsing a sentence:

sentence: Further modifications could make the project economically attractive.

phrase: X:agent make Y:object

What is the manifested discrepancy? Theoretically, a discrepancy is detected since the sentence includes 4 clauses while the phrase anticipates only 3 clauses. The words *economically attractive* seem extraneous. However, since the parser strives to justify the sentence by all syntactic means, it uses the words *economically attractive* as an adjunct, and so the problem actually goes undetected. Thus, how can a parsing program detect lexical gaps in the presence of multiple possible parses? Although we have developed a number of heuristic techniques this general problem is yet unresolved.

3.2 Clustering Phrases hy Text Examples

Can phrase properties be acquired from corpus by a clustering algorithm? A clustering algorithm such as [Michalsky and Stepp, 1982] requires two distinct sets of positive (IN) and negative (OUT) examples. Unfortunately, IN and OUT examples are not separated in the corpus. Consider the following example sentences:

(9) It can be made final after
a 45-day comment period.

(10) PTE will sell equipment it now
makes available by leasing.

(11) A tax credit of $13.5 million
made final net 87 million.

(12) The agency said that would
make it the most expensive cleanup.

Which of (9)-(12) are instances of Phrase2 (i.e., *he made it attractive)?* The positive, IN examples are (9,10,12). However they are not marked as such and so they are difficult to identify. Fortunately, general grammar rules can be applied to resolve grammatic variants. (9) is identified (appropriately) as a passive-variant of Phrase2. However, (10)-(12) are more problematic since no general grammar rule applies. (10) is a lexical variant of phrase2, but it is not recognized as such. (11) is not a lexical variant although a parse exists to justify that incorrect assumption (i.e., X made <final><net 87 million>). (12) is a variant of Phrase2 (NP replaces AP) but it goes undetected. Accordingly, by pursuing structure and surface similarities, a clustering algorithm places (9) and (11) in the same cluster (due to the word "final"); it fails in placing together (9), (10), and (12).

Thus, how can text examples be used in acquisition if IN examples cannot be identified? In our algorithm this problem is alleviated by the use of given lexical categories which predict for unknown phrases their possible variants.

3.3 Predicting Properties via Lexical Categories

Thus, from the corpus we cannot determine a phrase stamp (recall Section 3.2), nor phrase semantics (text itself provides no semantic clues. All one gets in text are words). Thus, in the absence of other available resources, can we capitalize on lexical categories?

Predicting Semantics: consider the following category (sensing verbs), identified by shared syntax and semantics:

(13) John saw Mary walking by.
(14) John heard Mary walking by.
(15) John sensed Mary walking by.

The properties shared by this category are given as follows:

- Shared syntax: VP — > V, NP, gerund-Sbar
- Shared semantics: object raising (the object of main phrase is the subject of the embedded phrase; the verb takes only a single object as a semantic argument).

Although this category seems quite consistent, in general consistency cannot be guaranteed and exceptions are pervasive [1].

Predicting Lexical Variants: consider the following category (dative verbs):

(16) John gave Mary a book.
(17) John sold Mary a book.
(18) John handed Mary a book.
BUT:
(19) *John donated Mary the book.

Donate presents an exception to the dative verb category [2], with regard to its allowed variants. *Donate's* stamp, it turns out, does not include the variant: donate(NP, NP, NP). Donate allows only one variant: donate(NP, NP, toNP).

Semantic bootstrapping, has been argued extensively [Grimshaw, 1981, Pinker, 1984, Levin, 1987, Kegl, 1987, Miller, 1985, Zernik, 1987] and its limitations are well established: lexical categories correlate syntactic and semantic features only to a limited extent. Thus, we seek a computational algorithm which can capitalize on lexical categories in spite of lexical exceptions.

The Algorithm: Indexing, Prediction, Validation

The algorithm is activated by a sentence presenting an unknown phrase:

(20) The agency said that would make
it the most expensive cleanup.

It proceeds in four steps:

Step 1 - Extract Argument Structure: from the given example identify the canonical lexical form. Bringing a phrase to a canonical form (identifying its argument structure) requires (a) accounting for general grammar transformations (e.g., passivization); (b) identifying lexical complements and factoring out adjuncts [Dyer and Zernik, 1986]. The obtained lexical canonical form for sentence (20) is given below:
subject:NP verb:make object:NP object:NP

Note that the main verb of the sentence *(the agency said)* is factored out.

Step 2 - Index a Category: through the given argument structure (i.e., NP, NP, NP) identify the matching categories. Five categories are indexed in this example:

dative:
John gave Mary a book.

beneficiary:
John baked Mary a cake.

selection:
Mexico elected Salinas president.

communication:
John told Mary a story.

state-change:
John painted the wall green.

Step 3 - Scan Corpus: each category has a stamp, a set of lexical variants. Consider the stamps of our five categories:

dative:
{(NP, NP1, NP2), (NP, NP2, to-NP1)}

beneficiary:
{(NP, NP1, NP2), (NP, NP2, for-NP1), (NP, NP2)}

selection:
{(NP, NP1, NP2), (NP, NP1, to-be-NP2)}

communication:
{(NP, NP1, NP2), (NP, NP1, Sbar)}

state-change:
{(NP, NP1, NP2), (NP, NP1, AP), (NP, AP, NP1)}

For each category check whether each variant is found in the corpus. Only a category whose entire stamp is found in the corpus is taken as a candidate. The obtained results are given below:

- Dative is rejected since the variant NP make NP to NP is not found in the corpus.
- Communication is rejected since the variant NP make NP Sbar is not found.
- Selection is rejected since the variant X make Y to be Z is not found.
- Beneficiary (inappropriately) and state-change (appropriately) are both completely validated.

Since both beneficiary and state-change are indexed, the entire process is not unambiguous.

Step 4 - Predict Semantics: through the selected categories determine the semantic structure of the phrase. In our example the categories of beneficiary and state-change determine the semantics of the new phrase:

PhraseA (beneficiary)	PhraseB (state-change)
act create	act state-change
actor X	cause X
object Z	object Y
beneficiary Y	new-attribute Z
	causing-act create

In fact, both phrases phrase A and phraseB are lexically correct. The correctness of PhraseA is manifested by sentence (21) below:

(21) John made Mary a meal.

[1]A problem totally ignored in this work is raised by sentences such as *John told Mary walking by is not a solution.* Here an inaccurate yet legitimate syntactic analysis, i.e. <<john told Mary walking by> is not a solution>, can induce incorrect semantic prediction.

[2]Indeed language learners make overgeneralization errors with promise and donate, as well as with other lexical exceptions.

However, only PhraseB provides a correct interpretation to sentence (20) above. In absence of reference resolution a parser cannot rule out that *make it an expensive cleanup* is not a beneficiary case.

5 Evaluation

We estimate the merits and the limitations of this new approach.

5.1 Qualitative Evaluation

Semantic approximation: with this method the semantics acquired for verbs present only an approximation to the intended meaning. Smell, hear, and taste, for example cannot be distinguished as to the sensor type. The general category provides for all those verbs only the aspect of a sensory state.

Errors in processing corpus: this method requires processed corpus. However, in the presence of lexical unknowns (this is always a realistic assumption) the given parse is not perfect. Two typical parsing errors are given below:

(22) Brazil will make <$350 million >< interest payment>.
(23) PTE made <final net><$2 million>.

Inaccurate parsing is problematic since there is no error indication.

Inapplicable phrases: this method is ineffective for a large set of phrases. A phrase with a simple argument structure (e.g., liquidate) cannot be processed since it does not index any particular category. Learning by analogy requires strong similar syntactic features.

Complex categories: complex categories such as *earn* and *raisin a* [Bresnan, 1982] are difficult to distinguish *(John wanted Mary to go* vs. *John expected Mary to go,* respectively) [Boguraev, 1988], since their syntax is identical. Using a large corpus, special variants are identified (i.e., *he waiits that Mary go)* which support discrimination.

5.2 Quantitative Evaluation

This algorithm was tested so far by four different phrases. The entire corpus included 15,000 sentences. For each phrase we first "grepped" (used the grep command) the relevant sentences. For MAKE, for example, we experimented with about 400 samples. Learning is relative to more than 100 lexical categories.

Clearly, the sample so far is too small to draw statistically significant conclusions. Eventually, we will gather three statistical factors.

Parsing evaluation: the number of parsing failures (the ratio before and after learning) is the overall measure of success. However, this is a tough figure to determine for two reasons: (1) in cases of ambiguity it is not always objectively clear what the appropriate interpretation should be; (2) it is not possible to automatically identify a parsing failure. Most parsing failures go undetected.

Applicable phrases: the merit of the method will eventually be determined by the number of phrases which could theoretically be acquired.

Correct/incorrect learning: among the set of applicable phrases how many are appropriately acquired? This factor is subject to refinement of representation and augmentation of parsing quality (of the corpus).

6 On the Availability of Resources

We examine the list of alternative knowledge sources which could potentially be exploited in lexicon acquisition.

Learn from text examples: Text examples are readily available on line. However, as shown in this paper, semantics cannot be extracted from surface examples only. Even the apparently simple task of identifying selectional restrictions is non-trivial:

(24) He placed Mozart on the shelf.
(25) He took it up with his dad.

Due to metonymy (24) and pronoun (25), a learning algorithm might induce incorrect selectional restrictions for the intended phrases.

Learn from a user: a user, frequently an in-house lexicographer, could encode phrases manually. However, experience indicates two principal problems: (a) systematic encoding cannot be imposed unless acquisition is mechanized; (b) exhaustive encoding is difficult to carry out manually. A human operator cannot exhaust from corpus all variants of a phrase. A computer program can outperform a lexicographer in exhausting a corpus and in discovering overlooked properties.

Learn from context: the association of a word and its meaning can best be accomplished by learning words in context [Anderson, 1981, Granger, 1977, Selfridge, 1986, Jacobs and Zernik, 1988]. A program is presented with a pair: the new word and a conceptual representation of the context. However, this entails one tough condition. The full conceptual context must be hand encoded for each learning episode. Encoding lexical knowledge directly turns out easier than encoding the entire context for each case.

Learn from on-line dictionary: existing dictionaries such as Longman's, Webster's and Roget's could be exploited [Byrd *et al*., 1988, Boguraev, 1988]. A word is defined by its (a) grammar code (i.e., subcategorization and lexical variants); (b) is-a and part-of definitions given by text examples; (c) verb semantics (i.e., what is the meaning of *expect* or *make?)* given also by text examples. Objects such as zoological concepts lend themselves to is-a and part-of organization. However, verb semantics defy such dichotomy. In general such verbs are described in the dictionary by free-form text examples, which in turn require learning from text.

Learn from lexical categories: in absence of other resources, human learners guess word properties by identifying similarities with other, well-known words [Miller, 1985]. However, this method is error prone and it must be augmented by evidence found in corpus. Lexical categories can project a hypothesis that can be validated or disproved by text examples.

7 Conclusions

In this work we investigate the possibility of exploiting on-line text for lexical acquisition. This enterprise is promising since text examples reveal overt as well as covert properties of lexical items such as verb phrases. However, while invaluable to human lexicographers, the process of scanning corpus does not lend itself readily to mechanical lexicography for one simple reason: text examples do not directly provide either syntax or semantics. Sentences must first be processed in order for lexical properties to be extracted.

Consequently, the major research dilemma regards pre-processing - the depth of sentential analysis required by the learning program. On the one hand, a preprocessor that provides deep syntactic and semantic analysis seems very useful in learning. However, such a processor is problematic for one practical and one theoretical reasons: first, no existing processor can provide in-depth (or even shallow) analysis of extensive corpus; second, due to ambiguity, a processor must commit itself to certain interpretations and thereoff might bias the learning process (as shown by inaccurate parsing examples throughout this paper).

On the other hand, raw text with minimal preprocessing is not useful since surface features (e.g., word co-occurrence) and simple clustering cannot contribute significant lexical features.

Therefore, we choose to pre-process sentences only to the extent that very general grammatic variants (e.g., passive voice) are factored out. We do not rely on lexical rules such as dative movement. In order to make predictions we use general lexical rules organized as lexical categories. Using this method we have shown that certain lexical properties such as rough semantics and phrase variants can be acquired successfully. In the future we intend to further investigate the interaction of pre-processing and learning.

References

[Anderson, 1981] J Anderson. *A Theory of Language Acquisition Based on General Learning Principles.* In *Proceedings Seventh International Joint Conference on Artificial Intelligence,* Vancouver, Canada, August, 1981.

[Boguraev, 1988] B Boguraev and T Briscoe. *Large Lexicons for Natural Language Processing.* Computational Linguistics, Vol 13 No 3-4, 1988.

[Bresnan, 1982] J Bresnan *Control and Complementation.* In *The Mental Representation of Grammatical Relations,* J Bresnan, editor. MIT Press, Cambridge, MA, 1982.

[Byrd *et al,* 1988] R Byrd, N Calzolari, M Chodorow, J Klavans, M Neff, and 0 Rizk. *Tools and Methods for Computational Linguistics.* Computational Linguistics, Vol 13 No 3-4, 1988.

[Dyer and Zernik, 1986] M Dyer and U Zernik. *Encoding and Acquiring Figurative Phrases in the Phrasal Lexicon.* In *Proceedings 24th Annual Meeting of the Association for Computational Linguistics,* New York, NY, 1986.

[Granger, 1977] R Granger. *FOULUP: A Program That Figures Out Meanings of Words from Context.* In *Proceedings Fifth International Joint Conference on Artificial Intelligence,* Cambridge, MA, August, 1977.

[Grimshaw, 1981] J Grimshaw. *Form, Function, and the Language Acquisition Device.* In *The Logical Problem of Language Acquisition,* C Baker and J McCarthy, editors. MIT Press, Cambridge, MA, 1981.

[Jacobs and Rau, 1988] P Jacobs and L Rau. *A Friendly Merger of Conceptual Expectations and Linguistic Analysis in a Text Processing System.* In *Proceedings of the Fourth Conference on Artificial Intelligence Applications,* San Diego, CA, March 1988.

[Jacobs and Zerni k, 1988] P Jacobs and U Zernik. *Learning Phrases from Text: A Case Study.* In *Proceedings of the Seventh National Conference on Artificial Intelligence,* St. Paul, Minesotta, August, 1988.

[Kegl, 1987] J Kegl. *Machine-Readable Dictionanes ana Education.* In *Lexicon Workshop: LEX,* D Walker, editor. CSLI, Palo Alto, CA, 1987.

[Levin, 1987] B Levin. *Approaches to Lexical Semantic Representation.* In *Lexicon Workshop: LEX,* D Walker, editor. CSLI, Palo Alto, CA, 1987.

[Marquez, 1986] G Marquez. *Life in the Time of Cholera.* Knopf, New York, NY, 1986.

[Michalsky and Stepp, 1982] R Michalsky and R Stepp. *Learning from Observation: Conceptual Clustering.* In *Machine Learning,* R Michalsky, J Carbonell, and T Mitchell, editors. Tioga Press, Palo Alto, CA, 1982.

[Miller, 1985] G Miller. Dictionaries of the Mind. In *Proceedings of the Twenty Third Annual Meeting of the Association for Computational Linguistics, CA*cago, IL, July 1985.

[Pinker, 1984] S Pinker. *Language Learnability and Language Development.* Harvard University Press, Cambridge, MA, 1984.

[Selfridge, 1986] M Selfridge. *A Computer Model of Child Language Learning.* Artificial Intelligence, Vol 29, 1986.

[Zernik, 1987] U Zernik. *Learning Phrases in a Hierarchy.* In *Proceedings 10th International Joint Conference on Artificial Intelligence,Milano* Italy, August 1987.