# Morphological Cues for Lexical Semantics

**Marc Light**
Seminar für Sprachwissenschaft
Universität Tübingen
Wilhelmstr. 113
D-72074 Tübingen
Germany
light@sfs.nphil.uni-tuebingen.de

## Abstract

Most natural language processing tasks require lexical semantic information. Automated acquisition of this information would thus increase the robustness and portability of NLP systems. This paper describes an acquisition method which makes use of fixed correspondences between derivational affixes and lexical semantic information. One advantage of this method, and of other methods that rely only on surface characteristics of language, is that the necessary input is currently available.

## 1 Introduction

Some natural language processing (NLP) tasks can be performed with only coarse-grained semantic information about individual words. For example, a system could utilize word frequency and a word cooccurrence matrix in order to perform information retrieval. However, many NLP tasks require at least a partial understanding of every sentence or utterance in the input and thus have a much greater need for lexical semantics. Natural language generation, providing a natural language front end to a database, information extraction, machine translation, and task-oriented dialogue understanding all require lexical semantics. The lexical semantic information commonly utilized includes verbal argument structure and selectional restrictions, corresponding nominal semantic class, verbal aspectual class, synonym and antonym relationships between words, and various verbal semantic features such as causation and manner.

Machine readable dictionaries do not include much of this information and it is difficult and time consuming to encode it by hand. As a consequence, current NLP systems have only small lexicons and thus can only operate in restricted domains. Automated methods for acquiring lexical semantics could increase both the robustness and the portability of such systems. In addition, such methods might provide insight into human language acquisition.

After considering different possible approaches to acquiring lexical semantic information, this paper concludes that a "surface cueing" approach is currently the most promising. It then introduces morphological cueing, a type of surface cueing, and discusses an implementation. It concludes by evaluating morphological cues with respect to a list of desiderata for good surface cues.

## 2 Approaches to Acquiring Lexical Semantics

One intuitively appealing idea is that humans acquire the meanings of words by relating them to semantic representations resulting from perceptual or cognitive processing. For example, in a situation where the father says *Kim is throwing the ball* and points at Kim who is throwing the ball, a child might be able learn what *throw* and *ball* mean. In the human language acquisition literature, Grimshaw (1981) and Pinker (1989) advocate this approach; others have described partial computer implementations: Pustejovsky (1988) and Siskind (1990). However, this approach cannot yet provide for the automatic acquisition of lexical semantics for use in NLP systems, because the input required must be hand coded: no current artificial intelligence system has the perceptual and cognitive capabilities required to produce the needed semantic representations.

Another approach would be to use the semantics of surrounding words in an utterance to constrain the meaning of an unknown word. Borrowing an example from Pinker (1994), upon hearing *I glipped the paper to shreds*, one could guess that the meaning of *glib* has something to do with tearing. Similarly, one could guess that *filp* means something like *eat* upon hearing *I filped the delicious sandwich and now I'm full.* These guesses are cued by the meanings of *paper, shreds, sandwich, delicious, full,* and the partial syntactic analysis of the utterances that contain them. Granger (1977), Berwick (1983), and Hastings (1994) describe computational systems that implement this approach. However, this approach is hindered by the need for a large amount of initial lexical semantic information and the need for a robust natural language understanding system that produces semantic representations as output, since producing this output requires precisely the lexical semantic information the system is trying to acquire.

A third approach does not require any semantic information related to perceptual input or the input utterance. Instead it makes use of fixed correspondences between surface characteristics of language input and lexical semantic information: surface characteristics serve as cues for lexical semantics of the words. For example, if a verb is seen with a noun phrase subject and a sentential complement, it often has verbal semantics involving spatial perception and cognition, *e.g.*, *believe, think, worry*, and *see* (Fisher, Gleitman, and Gleitman, 1991; Gleitman, 1990). Similarly, the occurrence of a verb in the progressive tense can be used as a cue for the non-stativeness of the verb (Dorr and Lee, 1992); stative verbs cannot appear in the progress tense (e.g.,*Mary is loving her new shoes). Another example is the use of patterns such as $NP, NP *$ *,and other*$NP$ to find lexical semantic information such as hyponym (Hearst, 1992). *Temples, treasuries, and other important civic buildings* is an example of this pattern and from it the information that temples and treasuries are types of civic buildings would be cued. Finally, inducing lexical semantics from distributional data (*e.g.*, (Brown et al., 1992; Church et al., 1989)) is also a form of surface cueing. It should be noted that the set of fixed correspondences between surface characteristics and lexical semantic information, at this point, have to be acquired through the analysis of the researcher—the issue of how the fixed correspondences can be automatically acquired will not be addressed here.

The main advantage of the surface cueing approach is that the input required is currently available: there is an ever increasing supply of online text, which can be automatically part-of-speech tagged, assigned shallow syntactic structure by robust partial parsing systems, and morphologically analyzed, all without any prior lexical semantics.

A possible disadvantage of surface cueing is that surface cues for a particular piece of lexical semantics might be difficult to uncover or they might not exist at all. In addition, the cues might not be present for the words of interest. Thus, it is an empirical question whether easily identifiable abundant surface cues exist for the needed lexical semantic information. The next section explores the possibility of using derivational affixes as surface cues for lexical semantics.

## 3 Morphological Cues for Lexical Semantic Information

Many derivational affixes only apply to bases with certain semantic characteristics and only produce derived forms with certain semantic characteristics. For example, the verbal prefix *un-* applies to telic verbs and produces telic derived forms. Thus, it is possible to use *un-* as a cue for telicity. By searching a sufficiently large corpus we should be able to identify a number of telic verbs. Examples from the Brown corpus include *clasp, coil, fasten, lace*, and *screw*.

A more implementation-oriented description of the process is the following: (i) analyze affixes by hand to gain fixed correspondences between affix and lexical semantic information (ii) collect a large corpus of text, (iii) tag it with part-of-speech tags, (iv) morphologically analyze its words, (v) assign word senses to the base and the derived forms of these analyses, and (vi) use this morphological structure plus fixed correspondences to assign semantics to both the base senses and the derived form senses. Step (i) amounts to doing a semantic analysis of a number of affixes the goal of which is to find semantic generalizations for an affix that hold for a large percentage of its instances. Finding the right generalizations and stating them explicitly can be time consuming but is only performed once. Tagging the corpus is necessary to make word sense disambiguation and morphological analysis easier. Word sense disambiguation is necessary because one needs to know which sense of the base is involved in a particular derived form, more specifically, to which sense should one assign the feature cued by the affix. For example, *stress* can be either a noun *the stress on the third syllable* or a verb *the advisor stressed the importance of finishing quickly.* Since the suffix *-ful* applies to nominal bases, only a noun reading is possible as the stem of *stressful* and thus one would attach the lexical semantics cued by *-ful* to the noun sense. However, *stress* has multiple readings even as a noun: it also has the reading exemplified by *the new parent was under a lot of stress.* Only this reading is possible for *stressful.*

In order to produce the results presented in the next section, the above steps were performed as follows. A set of 18 affixes were analyzed by hand providing the fixed correspondences between cue and semantics. The cued lexical semantic information was axiomatized using Episodic Logic (Hwang and Schubert, 1993), a situation-based extension of standard first order logic. The Penn Treebank version of the Brown corpus (Marcus, Santorini, and Marcinkiewicz, 1993) served as the corpus. Only its words and part-of-speech tags were utilized. Although these tags were corrected by hand, part-of-speech tagging can be automatically performed with an error rate of 3 to 4 percent (Merialdo, 1994; Brill, 1994). The Alvey morphological analyzer (Ritchie et al., 1992) was used to assign morphological structure. It uses a lexicon with just over 62,000 entries. This lexicon was derived from a machine readable dictionary but contains no semantic information. Word sense disambiguation for the bases and derived forms that could not be resolved using part-of-speech tags was not performed. However, there exist systems for such word sense disambiguation which do not require explicit lexical semantic information (Yarowsky, 1993; Schütze, 1992).

Let us consider an example. One sense of the suffix *-ize* applies to adjectival bases (*e.g., centralize*). This sense of the affix will be referred to as *-Aize*. (A related but different sense applies to nouns, *e.g., glamorize*. The part-of-speech of the base is used to disambiguate these two senses of *-ize*.) First, the regular expressions "`.*IZ(E|ING|ES|ED)$`" and "`^V.*`" are used to collect tokens from the corpus that were likely to have been derived using *-ize*. The Alvey morphological analyzer is then applied to each type. It strips off *-Aize* from a word if it can find an entry with a reference form of the appropriate orthographic shape and has the features "uninflected," "latinate," and "adjective." It may also build an appropriate base using other affixes, *e.g.*,[[*tradition -al*] *-Aize*].[1] Finally, all derived forms are assigned the lexical semantic feature CHANGE-OF-STATE and all the bases are assigned the lexical semantic feature IZE-DEPENDENT. Only the CHANGE-OF-STATE feature will be discussed here. It is defined by the axiom below.

```
For all predicates P with features
CHANGE-OF-STATE and DYADIC:
∀x,y,e [P(x,y)**e ->
        [∃e1:[at-end-of(e1,e) ∧
              cause(e,e1)]
         [rstate(P)(y)**e1] ∧
         ∃e2:at-beginning-of(e2,e)
         [¬rstate(P)(y)**e2]]]
```

The operator ** is analogous to $\models$ in situation semantics; it indicates, among other things, that a formula describes an event. P is a place holder for the semantic predicate corresponding to the word sense which has the feature. It is assumed that each word sense corresponds to a single semantic predicate. The axiom states that if a CHANGE-OF-STATE predicate describes an event, then the result state of this predicate holds at the end of this event and that it did not hold at the beginning, *e.g.*, if one wants to formalize something it must be non-formal to begin with and will be formal after. The result state of an *-Aize* predicate is the predicate corresponding to its base; this is stated in another axiom.

Precision figures for the method were collected as follows. The method returns a set of normalized (*i.e.*, uninflected) word/feature pairs. A human then determines which pairs are "correct" where correct means that the axiom defining the feature holds for the instances (tokens) of the word (type). Because of the lack of word senses, the semantics assigned to a particular word is only considered correct, if it holds for all senses occurring in the relevant derived word tokens.[2] For example, the axiom above must hold for all senses of *centralize* occurring in the corpus in order for the *centralize*/CHANGE-OF-STATE pair to be correct. The axiom for IZE-DEPENDENT must hold only for those senses of *central* that occur in the tokens of *centralize* for the *central*/IZE-DEPENDENT pair to be correct. This definition of correct was constructed, in part, to make relatively quick human judgements possible. It should also be noted that the semantic judgements require that the semantics be expressed in a precise way. This discipline is enforced in part by requiring that the features be axiomatized in a denotational logic. Another argument for such an axiomatization is that many NLP systems utilize a denotational logic for representing semantic information and thus the axioms provide a straightforward interface to the lexicon.

To return to our example, as shown in Table 1, there were 63 *-Aize* derived words (types) of which 78 percent conform to the CHANGE-OF-STATE axiom. Of the bases, 80 percent conform to the IZE-DEPENDENT axiom which will be discussed in the next section. Among the conforming words were *equalize, stabilize,* and *federalize.* Two words that seem to be derived using the *-ize* suffix but do not conform to the CHANGE-OF-STATE axiom are *penalize* and *socialize (with the guests)*. A different sort of non-conformity is produced when the morphological analyzer finds a spurious parse. For example, it analyzed *subsidize* as *[sub- [side -ize]]* and thus produced the *sidize*/CHANGE-OF-STATE pair which for the relevant tokens was incorrect. In the first sort, the non-conformity arises because the cue does not always correspond to the relevant lexical semantic information. In the second sort, the non-conformity arises because a cue has been found where one does not exist. A system that utilizes a lexicon so constructed is interested primarily in the overall precision of the information contained within and thus the results presented in the next section conflate these two types of false positives.

[1] In an alternative version of the method, the morphological analyzer is also able to construct a base on its own when it is unable to find an appropriate base in its lexicon. However, these "new" bases seldom correspond to actual words and thus the results presented here were derived using a morphological analyzer configured to only use bases that are directly in its lexicon or can be constructed from words in its lexicon.

[2] Although this definition is required for many cases, in the vast majority of the cases, the derived form and its base have only one possible sense (*e.g., stressful*).

## 4 Results

This section starts by discussing the semantics of 18 derivational affixes: *re-, un-, de-, -ize, -en, -ify, -le, -ate, -ee, -er, -ant, -age, -ment, mis-, -able, -ful, -less*, and *-ness*. Following this discussion, a table of precision statistics for the performance of these surface cues is presented. Due to space limitations, the lexical semantics cued by these affixes can only be loosely specified. However, they have been axiomatized in a fashion exemplified by the CHANGE-OF-STATE axiom above (see (Light, 1996; Light, 1992)).

The verbal prefixes *un-*, *de-*, and *re-* cue aspectual information for their base and derived forms. Some examples from the Brown corpus are *unfasten, unwind, decompose, defocus, reactivate*, and *readapt*. Above it was noted that *un-* is a cue for telicity. In fact, both *un-* and *de-* cue the CHANGE-OF-STATE feature for their base and derived forms—the CHANGE-OF-STATE feature entails the TELIC feature. In addition, for *un-* and *de-*, the result state of the derived form is the negation of the result state of the base (NEG-OF-BASE-IS-RSTATE), *e.g.*, the result of unfastening something is the opposite of the result of fastening it. As shown by examples like *reswim the last lap*, *re-* only cues the TELIC feature for its base and derived forms: the lap might have been swum previously and thus the negation of the result state does not have to have held previously (Dowty, 1979). For *re-*, the result state of the derived form is the same as that of the base (RSTATE-EQ-BASE-RSTATE), *e.g.*, the result of reactivating something is the same as activating it. In fact, if one reactivates something then it is also being activated: the derived form entails the base (ENTAILS-BASE). Finally, for *re-*, the derived form entails that its result state held previously, *e.g.*, if one recentralizes something then it must have been central at some point previous to the event of recentralization (PRESUPS-RSTATE).

The suffixes *-Aize, -Nize, -en, -Aify, -Nify* all cue the CHANGE-OF-STATE feature for their derived form as was discussed for *-Aize* above. Some exemplars are *centralize, formalize, categorize, colonize, brighten, stiffen, falsify, intensify, mummify*, and *glorify*. For *-Aize, -en* and *-Aify* a bit more can be said about the result state: it is the base predicate (RSTATE-EQ-BASE), *e.g.*, the result of formalizing something is that it is formal. Finally *-Aize, -en*, and *-Aify* cue the following feature for their bases: if a state holds of some individual then either an event described by the derived form predicate occurred previously or the predicate was always true of the individual (IZE-DEPENDENT), *e.g.*, if something is central then either it was centralized or it was always central.

The "suffixes" *-le* and *-ate* should really be called verbal endings since they are not suffixes in English, *i.e.*, if one strips them off one is seldom left with a word. (Consequently, only regular expressions were used to collect types; the morphological analyzer was not used.) Nonetheless, they cue lexical semantics and are easily identified. Some examples are *chuckle, dangle, alleviate*, and *assimilate*. The ending *-ate* cues a CHANGE-OF-STATE verb and *-le* an ACTIVITY verb.

The derived forms produced by *-ee*, *-er*, and *-ant* all refer to participants of an event described by their base (PART-IN-E). Some examples are *appointee, deportee, blower, campaigner, assailant*, and *claimant*. In addition, the derived form of *-ee* is also sentient of this event and non-volitional with respect to it (Barker, 1995).

The nominalizing suffixes *-age* and *-ment* both produce derived forms that refer to something resulting from an event of the verbal base predicate. Some examples are *blockage, seepage, marriage, payment, restatement, shipment*, and *treatment*. The derived forms of *-age* entail that an event occurred and refer to something resulting from it (EVENT-AND-RESULTANT)), *e.g.*, seepage entails that seeping took place and that the seepage resulted from this seeping. Similarly, the derived forms of *-ment* entail that an event took place and refer either to this event, the proposition that the event occurred, or something resulting from the event (REFERS-TO-E-OR-PROP-OR-RESULT), *e.g.*, a restatement entails that a restating occurred and refers either to this event, the proposition that the event occurred, or to the actual utterance or written document resulting from the restating event. (This analysis is based on (Zucchi, 1989).)

The verbal prefix *mis-*, *e.g.*, *miscalculate* and *misquote*, cues the feature that an action is performed in an incorrect manner (INCORRECT-MANNER). The suffix *-able* cues a feature that it is possible to perform some action (ABLE-TO-BE-PERFORMED), *e.g.*, something is enforceable if it is possible that something can enforce it (Dowty, 1979). The words derived using *-ness* refer to a state of something having the property of the base (STATE-OF-HAVING-PROP-OF-BASE), *e.g.*, in *Kim's fierceness at the meeting yesterday was unusual* the word *fierceness* refers to a state of Kim being fierce. The suffix *-ful* marks its base as abstract (ABSTRACT): *careful, peaceful, powerful*, etc. In addition, it marks its derived form as the antonym of a form derived by *-less* if it exists (LESS-ANTONYM). The suffix *-less* marks its derived forms with the analogous feature (FUL-ANTONYM). Some examples are *colorful/less, fearful/less, harmful/less*, and *tasteful/less*.

The precision statistics for the individual lexical semantic features discussed above are presented in Table 1 and Table 2. Lexical semantic information was collected for 2535 words (bases and derived forms). One way to summarize these tables is to calculate a single precision number for all the features in a table, *i.e.*, average the number of correct types for each affix, sum these averages, and then divide this sum by the total number of types. Using this statistic it can be said that if a random word is derived, its features have a 76 percent chance of being true and if it is a stem of a derived form, its features have a 82 percent chance of being true.

Computing recall requires finding all true tokens of a cue. This is a labor intensive task. It was performed for the verbal prefix *re-* and the recall was found to be 85 percent. The majority of the missed *re-* verbs were due to the fact that the system only looked at verbs starting with RE and not other parts-of-speech, *e.g.*, many nominalizations such as *reaccommodation* contain the *re-* morphological cue. However, increasing recall by looking at all open class categories would probably decrease precision. Another cause of reduced recall is that some stems were not in the Alvey lexicon or could not be properly extracted by the morphological analyzer. For example, *-Nize* could not be stripped from *hypothesize* because Alvey failed to reconstruct *hypothesis* from *hypothes*. However, for the affixes discussed here, 89 percent of the bases were present in the Alvey lexicon.

## 5 Evaluation

Good surface cues are easy to identify, abundant, and correspond to the needed lexical semantic information (Hearst (1992) identifies a similar set of desiderata). With respect to these desiderata, derivational morphology is both a good cue and a bad cue.

Let us start with why it is a bad cue: there may be no derivational cues for the lexical semantics of a particular word. This is not the case for other surface cues, *e.g.*, distributional cues exist for every word in a corpus. In addition, even if a derivational cue does exist, the reliability (on average approximately 76 percent) of the lexical semantic information is too low for many NLP tasks. This unreliability is due in part to the inherent exceptionality of lexical generalization and thus can be improved only partially.

However, derivational morphology is a good cue in the following ways. It provides exactly the type of lexical semantics needed for many NLP tasks: the affixes discussed in the previous section cued nominal semantic class, verbal aspectual class, antonym relationships between words, sentience, etc. In addition, working with the Brown corpus (1.1 million words) and 18 affixes provided such information for over 2500 words. Since corpora with over 40 million words are common and English has over 40 common derivational affixes, one would expect to be able to increase this number by an order of magnitude. In addition, most English words are either derived themselves or serve as bases of at least one derivational affix.[3] Finally, for some NLP tasks, 76 per-

| Feature | Affix | Types | Precision |
|---|---|---|---|
| TELIC | *re-* | 164 | 91% |
| RSTATE-EQ-BASE-RSTATE | *re-* | 164 | 65% |
| ENTAILS-BASE | *re-* | 164 | 65% |
| PRESUPS-RSTATE | *re-* | 164 | 65% |
| CHANGE-OF-STATE | *un-* | 23 | 100% |
| NEG-OF-BASE-IS-RSTATE | *un-* | 23 | 91% |
| CHANGE-OF-STATE | *de-* | 35 | 34% |
| NEG-OF-BASE-IS-RSTATE | *de-* | 35 | 20% |
| CHANGE-OF-STATE | *-Aize* | 63 | 78% |
| RSTATE-EQ-BASE | *-Aize* | 63 | 75% |
| CHANGE-OF-STATE | *-Nize* | 86 | 56% |
| ACTIVITY | *-le* | 71 | 55% |
| CHANGE-OF-STATE | *-en* | 36 | 100% |
| RSTATE-EQ-BASE | *-en* | 36 | 97% |
| CHANGE-OF-STATE | *-Aify* | 17 | 94% |
| RSTATE-EQ-BASE | *-Aify* | 17 | 58% |
| CHANGE-OF-STATE | *-Nify* | 21 | 67% |
| CHANGE-OF-STATE | *-ate* | 365 | 48% |
| PART-IN-E | *-ee* | 22 | 91% |
| SENTIENT | *-ee* | 22 | 82% |
| NON-VOLITIONAL | *-ee* | 22 | 68% |
| PART-IN-E | *-er* | 471 | 85% |
| PART-IN-E | *-ant* | 21 | 81% |
| EVENT-AND-RESULTANT | *-age* | 43 | 58% |
| REFERS-TO-E-OR-PROP-OR-RESULTANT | *-ment* | 166 | 88% |
| INCORRECT-MANNER | *mis-* | 21 | 86% |
| ABLE-TO-BE-PERFORMED | *-able* | 148 | 84% |
| STATE-OF-HAVING-PROP-OF-BASE | *-ness* | 307 | 97% |
| FUL-ANTONYM | *-less* | 22 | 77% |
| LESS-ANTONYM | *-ful* | 22 | 77% |

Table 1: Derived words

| Feature | Affix | Types | Precision |
|---|---|---|---|
| TELIC | *re-* | 164 | 91% |
| CHANGE-OF-STATE | *Vun-* | 23 | 91% |
| CHANGE-OF-STATE | *Vde-* | 33 | 36% |
| IZE-DEPENDENT | *-Aize* | 64 | 80% |
| IZE-DEPENDENT | *-en* | 36 | 72% |
| IZE-DEPENDENT | *-Aify* | 15 | 40% |
| ABSTRACT | *-ful* | 76 | 93% |

Table 2: Base words

[3] The following experiment supports this claim. Just

cent reliability may be adequate. In addition, some affixes are much more reliable cues than others and thus if higher reliability is required then only the affixes with high precision might be used.

The above discussion makes it clear that morphological cueing provides only a partial solution to the problem of acquiring lexical semantic information. However, as mentioned in section 2 there are many types of surface cues which correspond to a variety of lexical semantic information. A combination of cues should produce better precision where the same information is indicated by multiple cues. For example, the morphological cue *re-* indicates telicity and as mentioned above, the syntactic cue the progressive tense indicates non-stativity (Dorr and Lee, 1992). Since telicity is a type of non-stativity, the information is mutually supportive. In addition, using many different types of cues should provide a greater variety of information in general. Thus morphological cueing is best seen as one type of surface cueing that can be used in combination with others to provide lexical semantic information.

## 6 Acknowledgements

A portion of this work was performed at the University of Rochester Computer Science Department and supported by ONR/ARPA research grant number N00014-92-J-1512.

## References

Barker, Chris. 1995. The semantics of *-ee*. In *Proceedings of the SALT conference.*

Berwick, Robert. 1983. Learning word meanings from examples. In *Proceedings of the 8th International Joint Conference on Artificial Intelligence (IJCAI-83).*

Brill, Eric. 1994. Some advances in transformation-based part of speech tagging. In *Proceedings of the Twelfth National conference on Artificial Intelligence: American Association for Artificial Intelligence (AAAI).*

Brown, Peter F., Vincent J. Della Pietra, Peter V. deSouza, Jennifer C. Lai, and Robert L. Mercer. 1992. Class-based $n$-gram models of natural language. *Computational Linguistics*, 18(4).

Church, Kenneth, William Gale, Patrick Hanks, and Donald Hindle. 1989. Parsing, word associations and typical predicate-argument relations. In *International Workshop on Parsing Technologies*, pages 389–98.

Dorr, Bonnie J. and Ki Lee. 1992. Building a lexicon for machine translation: Use of corpora for aspectual classification of verbs. Technical Report CS-TR-2876, University of Maryland.

Dowty, David. 1979. *Word Meaning and Montague Grammar.* Reidel.

Fisher, Cynthia, Henry Gleitman, and Lila R. Gleitman. 1991. On the semantic content of subcategorization frames. *Cognitive Psychology*, 23(3):331–392.

Gleitman, Lila. 1990. The structural sources of verb meanings. *Language Acquisition*, 1:3–55.

Granger, R. 1977. Foulup: a program that figures out meanings of words from context. In *Proceedings of the 5th International Joint Conference on Artificial Intelligence.*

Grimshaw, Jane. 1981. Form, function, and the language acquisition device. In C. L. Baker and J. J. McCarthy, editors, *the logical problem of language acquisition.* MIT Press.

Hastings, Peter. 1994. *Automatic Acquistion of Word Meaning from Context.* Ph.D. thesis, University of Michigan.

Hearst, Marti. 1992. Automatic acquisition of hyponyms from large text corpora. In *Proceedings of the fifteenth International Conference on Computational Linguistics (COLING).*

Hwang, Chung Hee and Lenhart Schubert. 1993. Episodic logic: a comprehensive natural representation for language understanding. *Mind and Machine*, 3(4):381–419.

Light, Marc. 1992. Rehashing *Re-*. In *Proceedings of the Eastern States Conference on Linguistics.* Cornell University Linguistics Department Working Papers.

Light, Marc. 1996. *Morphological Cues for Lexical Semantics.* Ph.D. thesis, University of Rochester, Rochester, NY.

Marcus, Mitchell, Beatrice Santorini, and Mary Ann Marcinkiewicz. 1993. Building a large annotated corpus of English: The Penn Treebank. *Computational Linguistics*, 19(2):313–330.

Merialdo, Bernard. 1994. Tagging English text with a probabilistic model. *Computational Linguistics*, 20(2):155–172.

Pinker, Steven. 1989. *Learnability and Cognition: The Acquisition of Argument Structure.* MIT Press.

Pinker, Steven. 1994. How could a child use verb syntax to learn verb semantics? *Lingua*, 92:377–410.

Pustejovsky, James. 1988. Constraints on the acquisition of semantic knowledge. *International journal of intelligent systems*, 3:247–268.

over 400 open class words were picked randomly from the Brown corpus and the derived forms were marked by hand. Based on this data, a random open class word in the Brown corpus has a 17 percent chance of being derived, a 56 percent chance of being a stem of a derived form, and an 8 percent chance of being both.

Ritchie, Graeme D., Graham J. Russell, Alan W. Black, and Steve G. Pulman. 1992. *Computational Morphology: Practical Mechanisms for the English Lexicon.* MIT press.

Schütze, Hinrich. 1992. Word sense disambiguation with sublexical representations. In *Statistically-Based NLP Techniques (American Association for Artificial Intelligence Workshop, July 12-16, 1992, San Jose, CA.)*, pages 109–113.

Siskind, Jeffrey M. 1990. Acquiring core meanings of words, represented as Jackendoff-style conceptual structures, from correlated streams of linguistic and non-linguistic input. In *Proceedings of the 28th Meeting of the Association for Computational Linguistics.*

Yarowsky, David. 1993. One sense per collocation. In *Proceedings of the ARPA Workshop on Human Language Technology.* Morgan Kaufmann.

Zucchi, Alessandro. 1989. *The Language of Propositions and Events: Issues in the Syntax and the Semantics of Nominalization.* Ph.D. thesis, University of Massachusetts, Amherst, MA.