

Principled Disambiguation: Discriminating Adjective Senses with Modified Nouns

John S. Justeson*
SUNY Albany

Slava M. Katz†
IBM T. J. Watson Research Center

Recent corpus-based work on word sense disambiguation explores the application of statistical pattern recognition procedures to lexical co-occurrence data from very large text databases. In this paper we argue for a linguistically principled approach to disambiguation, in which relevant contextual clues are narrowly defined, in syntactic and semantic terms, and in which only highly reliable clues are exploited. Statistical methods play a definite role in this work, helping to organize and analyze data, but the disambiguation method itself does not employ statistical data or decision criteria. This approach results in improved understanding of the disambiguation problem both in general and on a word-specific basis and leads to broadly applicable and nearly errorless clues to word sense. The approach is illustrated by an experiment discriminating among the senses of adjectives, which have been relatively neglected in work on sense disambiguation. In particular, the paper assesses the potential of nouns for discriminating among the senses of adjectives that modify them. This assessment is based on an empirical study of five of the most frequent ambiguous adjectives in English: hard, light, old, right, and short. About three-quarters of all instances of these adjectives can be disambiguated almost errorlessly by the nouns they modify or by the syntactic constructions in which they occur. Such disambiguation requires only simple rules, which can be automated easily. Furthermore, a small number of semantic attributes supply a compact means of representing the noun clues in a very few rules. Clues other than nouns are required when modified nouns are not useable. The sense of an ambiguous modified noun may be needed to determine the relevant semantic attribute for disambiguation of a target adjective; and other adjectives, verbs, and grammatical constructions all show evidence of high reliability, and sometimes of high applicability, when they stand in specific, well-defined syntactic relations to the ambiguous adjective. Some of these clues, however, may be hard to automate.

1. Introduction

This paper is an application of corpus analysis to an issue in word sense disambiguation: to what extent adjectives' senses can be separated by the nouns they modify, or more generally, by the phrases they modify or take as complements. We are interested specifically in *principled* disambiguation—systematic interpretation using highly reliable inferences based on linguistically motivated features. The work is empirical, addressing five of the most common, broadly applicable adjectives in English: *hard*, *light*, *old*, *right*, and *short*.

Nouns are intrinsically suited for principled disambiguation of adjectives. Most adjectives normally designate an attribute of an entity designated by the noun or noun

* Department of Anthropology, State University of New York, Albany, NY 12222.

† Current address: Weston Language Research, 138 Weston Road, Weston, CT 06883.

phrase they modify. Often, one sense or group of senses designates an attribute that is much more typically relevant to the noun's referent. For example, when the adjective *old* characterizes a human being, as in *old man* or *she is old*, it usually means 'aged'—not 'used,' 'of long standing,' or 'former.' In such cases, the mutual relevance of the adjective and noun senses is content specific (semantic) rather than word specific (lexical). We presume that it is this semantic relation rather than a lexical association that is normally involved in disambiguation.¹ One of the issues to be addressed in disambiguation is the passage from word-specific evidence to conceptual representation, a problem that we do not pretend to have solved in this paper.

In addition, we consider the syntactic constructions within which adjectives occur and through which the phrases or clauses they modify are determined. For example, some adjectives have senses that are not used predicatively; the use of such an adjective as a predicate rules out such senses. The syntactic clues we recognize are as reliable as the noun clues.

Noun-based disambiguation of adjectives is of special interest because a selection among different attributes (adjective senses) is likely to be sensitive to the attribute bundles (noun senses) they characterize. It does turn out that a small number of semantic features of nouns do provide fairly high coverage and very high reliability in adjective sense discrimination. Noun-based disambiguation is a structured form of co-occurrence-based disambiguation, various forms of which are prominent in corpus-based work. In co-occurrence-based approaches, it is usual to take into account the entire set of words in the vicinity of a target (Maarek and Smadja 1989; Yarowsky 1992). Gale, Church, and Yarowsky (1992) demonstrate that high reliability and coverage are simultaneously attainable with such an approach. The type of work we are pursuing has the potential to be more readily interpretable, though it is more difficult to automate. We suspect that the success of the comparison of contexts in bulk is due in large part to the effect of a few highly structured types of clues, such as we examine, and in part to more diffuse clues of other types.

We use statistical inference methods as tools for analyzing and attempting to understand the problem of disambiguation and the potential of resources such as modified nouns for solving this problem. Methodologically, our statistical analyses and our extraction of disambiguating features from the corpus are straightforward, with one exception: in order to be able to base inference on already disambiguated subcorpora, we had to devise a way to adjust for the bias brought into the sample by the criteria through which they had been disambiguated; the required formulas are presented and illustrated in the Appendix.

The structure of the paper is as follows. In Section 2 we describe the notion of a *word sense indicator*; this defines the types of features we consider as clues for disambiguation generally, and how specifically we use nouns as clues to the recognition of adjective senses. In Section 3, we describe the type of data we use; by defining adjective senses in terms of the meanings of different antonyms, we can take advantage of a large database of examples for analysis, examples that are ideally disambiguated by the co-occurrence of adjectives and their antonyms in semantically concordant structures. Section 4 provides a detailed report on the structure and results of the investigation of noun-based disambiguation. Section 5 discusses more complex types of inference

¹ Word-specific relations between adjectives and nouns are idiomatic, non-compositional pairs (so-called "freezes") in which the adjective itself has no independent sense, e.g., *hard cash* and *short cut*. In some cases, such as *hard fact*, it is difficult to draw the line between a noun-specific sense (here, 'incontrovertible') and a compositional sense (e.g., 'inflexible, unyielding'); such indeterminacy is of course one of the sources of a freeze.

involved in some noun-based disambiguation and addresses the potential of other types of indicators for adjective senses.

2. Word Sense Indicators

Our problem is a specific case of the more general problem of finding clues within the context of a word that indicate its sense fairly reliably. Content words that have a close syntactic relation to one another are useful candidates for examination and are intuitively more likely to bear a close semantic relation than words that are near one another but are not related syntactically. One much-studied example is the semantic relation between a verb and its arguments (e.g., Boguraev et al. 1989; Church and Hanks 1989; Braden-Harder 1991; Hindle and Rooth 1991).

Discrimination among senses of adjectives based on the nouns they modify or of which they are predicated has been the subject of less intensive and systematic study. Determining the potential of this line of evidence is the focus of this paper. We do this by performing a noun-based disambiguation experiment. Certainly, some nouns are strongly associated with particular senses of some of the adjectives that modify them. This association can be illustrated for the ambiguous adjective *old*, which has senses roughly synonymous with *aged*, *long existing*, *former*, *used*, and *obsolete*, using sentences from our experimental corpus (see Section 4.1). Two of the nouns most frequently modified by *old* in general texts are *man* and *house*. Overwhelmingly, *old* is used in the sense ‘aged (not young)’ when it modifies *man*, e.g., in

The man was very old and very frail, a widower.

He was a strong old man: he had lived through forty-five years of those wretched casseroles, but she missed him already.

“Guilty!” came the hoarse croaking sounds of the old men.

In some sentences, in fact, this noun is the only real basis, within the sentence itself, for inferring the sense of *old*:

The old man answered this time.

“Leave the other to the old man.”

“All except the old man.”

Man, therefore, can be taken as a fairly good indicator of the ‘aged’ sense of *old*. Similarly, when *old* modifies *house*, it almost always has one of its ‘not new’ senses, as in

In the fashionable suburb of Kingston, full of beautiful old houses. . .

. . . around the old Holton house he made many improvements. . .

He saw the tractors come and tear down the old houses and plow up the land. . .

When *old* modifies *house*, then, this is a good indication that *old* is being used in one of these senses. So *man* and *house* are reliably associated with different senses of *old*.

We refer to *man* and *house* as *indicators* for the senses of *old*. More generally, a feature *F* (such as the modified noun *man*) that is associated with a target word *T*

(such as the modifying adjective *old*) is an indicator for a sense S_i (e.g., 'aged') of T if, when the feature is present, that sense S_i is more likely than the other senses S_j , $j \neq i$. This statement is formalized in the Appendix.

This characterization of the indicators for word senses also provides flexibility regarding sense definition. Not only does it not require a single assessment, once and for all, concerning what the senses of a word may be, it requires no commitment to the reality of word senses themselves, as classically construed (see Pustejovsky and Boguraev 1993). Disambiguation may be pursued relative to many distinct issues, e.g., grammatical class, functional role, document topic, or lexical translation equivalent; the entities to be discriminated are the effective "senses" being identified. In this paper, we disambiguate relative to a pair of word sense groups, operationally by disambiguating relative to sense-specific antonyms; an old man, for example, is a man who is not young, and an old house is a house that is not new. Such sense distinctions have some justification in terms of the semantic organization of adjectives (see Gross, Fischer, and Miller 1989). However, our purpose in choosing them was purely for convenience in designing an experiment useful for determining the potential of noun-based disambiguation of adjectives.

The characterization of indicators is equally flexible with regard to the domain of its own applicability. It need not be assumed that all instances of the target word T are to be included in assessing the relative probabilities of different senses. In the experiment described here, we are determining the extent to which those senses of an adjective that are associated with one antonym can be distinguished from those associated with a different antonym according to the nouns that the target adjective modifies. We therefore discriminate between just these two sets of senses, which constitute the great majority of instances of the targets, and we exclude from both investigation and evaluation all instances in which the sense of the target does not fall in one of these two groups. We exclude all freezes from consideration as not being legitimate instances in which the adjectives actually have a definable sense (see footnote 1). In addition, we exclude the minority of instances that have definable senses that do not fall within these two groups. For example, *short* has a sense 'inadequate' that is related historically to its dimensional senses; however, this sense does not have a lexically specific antonym, whereas the dimensional senses do (*long* and *tall*).

3. Disambiguated Subcorpora

To extract a reasonable number of nouns that are indicators for the senses of target adjectives, one straightforward approach would be to extract a representative sample of sentences for each target adjective, to disambiguate each target manually, and to extract those nouns that are relatively frequent and that are modified by the target in one sense but not in the other. We adopted a different strategy, one that provided us with a large set of sentences in which target adjectives could be disambiguated automatically and with complete reliability. This strategy involved disambiguation of adjectives by their co-occurrence with sense-specific antonyms.

Antonym co-occurrence is a frequent and pervasive phenomenon, and it takes place under highly restricted semantic and syntactic conditions (Charles and Miller 1989; Justeson and Katz 1991, 1992). An adjective and its antonym refer to opposed values of the same attribute. When they modify the same noun in a sentence—this is the usual case in sentences in which they both occur—this attribute is virtually ensured of applying in a consistent way to both instances of the noun. When an adjective like *old* has different senses that are associated with different antonyms (like *new* and *young*), the adjective in these sentences is disambiguated by its antonym. Thus, in sentences

in which *old* and *young* modify the same noun, e.g., *man*, *old* is thereby interpretable as 'not young'; in those in which *old* and *new* both modify a noun, e.g., *house*, *old* is thereby interpretable as 'not new.'

The reason for this effect is easy to observe. Antonyms most often co-occur in direct comparisons or in contrastive opposition, directly reflecting both the identity of the attribute to which they pertain and the contrast in its value. As illustrated by the following sentences from the Brown Corpus (Francis and Kucera 1982), they usually occur in otherwise essentially identical phrases:

Photograph shows the wrong side of work with light strand being picked up under dark strand in position to be purled.

They indicated that no new errors were being made and that all old errors would be corrected within 60 days.

Note how easy it is to find synonyms for the epithet "miser" and how hard to find synonyms for "spendthrift."

When her right hand was incapacitated by the rheumatism, Sadie learned to write with her left hand.

I found myself becoming one of that group of people who, in Carlyle's words, are forever gazing into their own navels, anxiously asking am I right, am I wrong?

Radio broadcasts, however—now that even plain people could afford loud speakers on their sets—held old fans to the major-league races and attracted new ones. . .

We often say of a person that he looks young for his age or old for his age.

We refer to this pattern as *phrasal substitution*. In these cases, the phrases involved usually stand in direct semantic opposition (e.g., *her right hand*, . . . *her left hand*). Co-occurring antonyms are also frequently joined by *and* or *or*, or appear in noun phrases joined by prepositions and having the same head noun:

It was pitiful to see the thin ranks of warriors, old and young. . .

That was one more reason she didn't look forward to Cathy's visit, short or long;

As for this rider, I never saw him before or afterwards and never saw him dismounted, so whether he stood tall or short in his shoes, I can't say;

Skin colors range from white to dark brown, heights from short to tall, hair from long and straight to short and tightly curled.

It was a winter world without details, a world of shapes in an expanse ranging in color from light to dark gray.

But there is no sudden transition from hard rock to soft.

The chief function of these conjoined and prepositional co-occurrences is to cover the range of possible values of an attribute.

Whether the antonym co-occurrences involve contrastive opposition or a range of attribute values, they call forth the semantic dimension designated by the antonym pairs and guarantee concordance in adjective sense of the co-occurring antonyms. Thus, when an adjective has different sense-specific antonyms, their co-occurrences

as modifiers of different instances of the same noun reliably disambiguate that adjective. Five common English adjectives have such antonyms, yielding ten antonymous adjective pairs: *hard–easy*, *hard–soft*; *light–dark*, *light–heavy*; *old–new*, *old–young*; *right–left*, *right–wrong*; and *short–long*, *short–tall*. All the example sentences above involve one of these ten pairs, and they exemplify the concordance of the antonyms' senses. Furthermore, certain departures from the perfect phrasal substitution patterns equally constrain the senses of the antonyms to be concordant. These departures include insertion of adverbial modifiers, substitution of other words besides the antonyms, or use of the antonyms to modify noun phrases having the same head noun in phrases that otherwise differ from one another. The subcorpora that were used in the work reported here consist of those sentences in which a target adjective and its antonym modify separate instances of the same noun or clause.

4. The Experiment

This section describes our investigation of noun-based disambiguation and its results. Subcorpora were extracted for each of the five target adjectives, consisting of sentences in which the target was disambiguated by its co-occurrence with an antonym as modifiers of the same noun (Section 4.1). The tendency toward sense specificity of nouns modified by target adjectives in these sentences is demonstrated by showing that there is little overlap in the set of nouns modified by the target in the two antonym co-occurrence subcorpora for that target (Section 4.2). Nouns indicating the antonym-specific senses of these targets were then extracted by statistical analysis of their sense preferences. It is shown that these indicator nouns are also specific to the senses of the target adjectives in the corpus at large by using them successfully in a disambiguation procedure applied to 500 randomly selected sentences (Section 4.3).

Simple, broadly applicable semantic features characterize most of the indicator nouns, whereas broadly applicable syntactic features characterize many of their contexts. Together, these features discriminate the target senses, permitting a more compact and conceptual rather than word-specific representation of the indicators (Section 4.4): about three-quarters of the adjective instances are disambiguated by these features, and virtually errorlessly.

4.1 Acquiring Disambiguated Subcorpora

Our original study of antonym co-occurrence (Justeson and Katz 1991) was based on a version of the Brown Corpus, containing 54,717 sentences; it yields only 57 sentences in which both adjectives from any of the ten antonym pairs analyzed in the present study were modifying different instances of the same noun. To get enough sentences containing antonym co-occurrences of antonyms to address disambiguation issues adequately, we used the 1.5 million-sentence APHB Corpus. This corpus of 25,000,000 words was obtained from the American Printing House for the Blind and archived at IBM's T.J. Watson Research Center. It consists of stories and articles from books and general circulation magazines.

All sentences containing co-occurrences of the target adjective and each of its antonyms were extracted from the APHB Corpus, yielding 4391 sentential co-occurrences. These sentences were manually postprocessed to eliminate all instances in which either the target or its antonym was not being used adjectivally.² From the

² This could have been automated using a parser. Our immediate interest, however, is in discovering actual patterns of usage and not in building an automatic system. We did the work manually to avoid

remaining sentences, we further extracted a subset of sentences in which both members of the pair modify distinct instances of the same noun. This yielded 1487 sentences in which at least one of the target adjectives co-occurs with one of its antonyms, with both the target and its antonym modifying instances of the same noun that are in separate phrases. Some of these sentences have more than one such co-occurrence, so these sentences yielded 1535 total co-occurrences. Every one of these co-occurrences had sense-concordant antonyms modifying the same noun; any other sense is usually semantically incongruous, especially in direct phrasal substitutions. The sentences in which co-occurring antonymous adjectives modify the same noun therefore constitute a subcorpus in which the ambiguous members of the antonym pairs are discriminated relative to their antonym-specific senses. This gave us ten subcorpora, one for each antonym pair, of 1535 examples to use as a database for studying the extent to which modified nouns disambiguate their modifying adjectives.

In counting instances of nouns associated with each adjective, elided nouns and anaphoric pronouns were resolved (manually) whenever possible, adding to the counts for the noun referent, since we are studying the phenomenon of the adjective–noun relation. In addition, we stripped morphological suffixes from noun phrases, to recover an adjective–noun base. Thus *right winger*, *right fielder*, and *heavy sleeper* are recognized as deriving from *right wing*, *right field*, and *heavy sleep*; *short-staffed*, *short-lived*, and *light industrial* are recognized as derived from *short staff*, *short life*, and *light industry*. It was counted directly only in non-anaphoric usage. The sentences of this subcorpus contain 1535 such co-occurrences of the target adjectives and their antonyms. (Co-occurrence counts for each of the ten antonym pairs are given in Table 3 of the Appendix.)

4.2 Specificity of Nouns for Adjective Senses in the Disambiguated Subcorpora

The disambiguated subcorpora can be used to assess the extent to which target adjectives, when modifying a given noun, are specific to a single sense rather than being usable in either sense. They are in fact surprisingly consistent in this regard, as can be illustrated for the indicator nouns discussed in Section 2. There it was noted that *man* is an indicator of the ‘aged’ sense of *old* (with antonym *young*) and *house* for some of the ‘not new’ senses of *old* (with antonym *new*).

The specific association of *man* with the ‘aged’ sense of *old* is reflected in the use of this noun in antonymic constructions. The APHB Corpus contains 64 sentences in which both *young* and *old* modify *man*; e.g.,

*In Bihzad’s paintings we see people and animals as individuals—rich **men** and poor **men, old and young**, the elders in the mosque and the herdsmen camping among their horses in the fields.*

***Old men** saw visions and **young men** dreamed dreams.*

*I am an **old man**, you a **young** one; . . .*

*The **old man** turns to the **young** one and says: “The time has come for a few questions.”*

*In both the **old** and the **young man** this was a breach of habit.*

Most of these sentences involve phrasal substitution patterns typical of antonyms generally. In contrast, there is not one sentence in which both *old* and *new* modify either *man* or *men*.

the systematic errors to which all automatic parsers are subject.

The converse result obtains in the case of *house*. Four sentences contain instances in which *old* and *new* both modify *house* or *houses*:

*Cast-iron balustrades became the fashion, to be sought out when **old houses** were pulled down and removed to **new houses**. . .*

*. . . entire crates of dishes have been smashed when the trailers cross railroad tracks or other rough spots located between the **old** and the **new house**.*

*Fireplaces in the **new house**, but not in the **old one**?*

*Section 235 (of the Housing Act of 1968) helps families with low and "moderate" incomes to buy one- or two-family **houses, old or new**.*

In contrast, the APHB Corpus contains no sentence in which both *old* and *young* modify *house* or *houses*.

The specificity of *man* and *house* to particular senses of *old* is typical of nouns in these subcorpora. Most nouns by far were modified by the target in only one of its senses, in our co-occurrence sentences. We demonstrated this sense specificity of modified nouns by compiling all pairs consisting of a target adjective modifying the same noun as either of its antonyms. For each of these adjective–noun pairs, we determined the number of instances involving each sense; we want to determine the extent to which a particular adjective–noun pair tends to occur with only one of the two senses. One hundred and eighty-one adjective–noun pairs occur more than once,³ with a total of 1096 occurrences. Most by far have $n : 0$ sense distributions, i.e., the pairs occur in only one of the target's senses. Sixty-one adjective–noun pairs, covering 828 instances, have 4 or more instances each, and thus could admit 2 or more instances in their minority sense. Only 4 do have so many minority instances, covering 39 of the 828 total occurrences. The occurrence of 2 or more minority instances is not mainly a frequency effect; these 4 adjective–noun pairings are no more frequent on average than those that do not. So adjective–noun pairs do, as a rule, strongly favor one particular sense, and this is as true of pairs with many instances as of those with few.

4.3 Adjective Disambiguation Using Indicator Nouns

The specificity of nouns in the disambiguated corpus for senses of the target adjectives suggests potentially very high reliability for a noun-based procedure to disambiguate common adjectives. We evaluate the potential of such a procedure by extracting, from the co-occurrence sentences, a set of nouns that are indicators for the senses of the target adjectives and applying them to instances of the targets from non-co-occurrence sentences in the corpus at large.

Because adjectives co-occur with their antonyms fairly frequently, it was practical to extract disambiguated subcorpora large enough to provide a base for statistical inference. However, subcorpora in which most sentences exhibit phrasal substitution of antonyms are clearly not representative samples of the use of the target adjectives. This raises the possibility that the specificity of nouns for target adjective senses might be influenced by the nature of the sentences in which they occur—those that contain largely repeated, contrastive structures; we need evidence concerning their sense specificity from the corpus at large. This unrepresentativeness also introduces bias into the

³ If a pair occurs only once, there is no opportunity for its target adjective to appear sometimes in one sense and sometimes in another; so such pairs cannot be used in estimating the consistency of sense selection by nouns.

statistical process of inferring sense indicators for the corpus at large from the specially selected subcorpora, a bias for which we must correct. The Appendix gives the formula needed to project, from the disambiguated subcorpora to the corpus at large, the probability of each sense of the target adjective given the noun it modifies. This adjustment for bias requires that we know the sense distribution of the target adjectives in the corpus to which the indicators are intended to apply and the number of instances in the subcorpora in which each potential indicator noun is modified by the target adjective in each of its senses.

We therefore extracted from the APHB Corpus a random sample of 100 sentences containing adjectival instances of each target adjective, for a total of 500 sentences in all. In addition to requiring that the target word functions as an adjective, we exclude all freezes,⁴ as well as quantificational expressions such as *three years old* in which the target labels an attribute (e.g., *age*) rather than a value of that attribute (e.g., 'aged').⁵ The target adjectives in all 500 test sentences were manually disambiguated, both with respect to the antonyms and in some cases with respect to other senses not associated with either antonym. The distribution of antonym-related senses is given in the third column of Table 1. Since we are interested in discriminating between the two antonym-related sets of senses of the targets, we limited attention to those instances of that target occurring in a sense for which an antonym exists. For this reason, the total number of instances is less than 100 for each target adjective, varying from 89 for *hard* to 99 for *right*. After determining the sense distribution for the target in each set, we could project which nouns in the subcorpora are likely to be sense indicators for the target adjectives in these samples.

The following nouns were projected to show a preference for one or the other sense of the target adjectives that was statistically significant at the .05 level (nouns are in roman, classes in italic):

hard-not easy: *it, clauses*

hard-not soft: *none*

light-not dark: *none*

light-not heavy: *cruiser, harness, load*

old-not new: *proper names (of things), world, thing, car, way*

old-not young: *man, people, proper names (of people), woman, proper names (of places), lady, you, wine, person, bull, he, I, one (animate pronoun)*

right-not left: *hand*

right-not wrong: *he, I, thing, way, what, clauses, answer, proper names (of people), act, country, decision, expert, masturbation, note, people, reason, technician, that, theory, you*

short-not long: *term, syllable, hair, range, run, story*

short-not tall: *none*

⁴ When an adjective in a frozen noun phrase contributes no distinct meaning to the phrase (see footnote 1), it is excluded from consideration; such freezes should be found in the dictionary, and their identification in text is a separate problem.

⁵ The total number of sentences containing the target word that were extracted to yield 100 such adjectival instances of the target was 160 for *hard*, 444 for *light*, 124 for *old*, 278 for *right*, and 147 for *short*.

Table 1

Coverage and disambiguation error rates for target adjectives in 100-sentence samples, using different indicator sets.

Sense Distribution in 100-Sentence Samples			Indicator Nouns		All Nouns		Syntax and Sem. Attributes	
			Coverage	Errors	Coverage	Errors	Coverage	Errors
Hard	Not easy	69	33	0	33	1	66	0
	Not soft	20	0	0	6	0	14	0
Light	Not dark	31	0	0	23	2	17	0
	Not heavy	67	0	0	8	0	19	0
Old	Not new	48	3	0	20	3	38	0
	Not young	44	28	0	36	0	32	2
Right	Not left	22	8	0	20	1	17	0
	Not wrong	77	43	0	47	0	67	0
Short	Not long	86	11	0	38	0	73	0
	Not tall	9	0	0	6	0	5	1
			126	0	237	7	348	3
Overall		473	26.6%	0%	50.1%	3.0%	73.6%	0.9%

The statistical procedure that was used to identify these nouns as indicators is described in detail in the Appendix. The number of significant indicators recovered is quite variable, ranging from none for the 'not soft' sense of *hard*, the 'not dark' sense of *light*, and the 'not tall' sense of *short*, to 13 for the 'not young' sense of *old* and 20 for the 'not wrong' sense of *right*.

At this point, we had extracted a small set of statistically significant nouns that are projected to be indicators for adjective senses in the random samples. We then identified each instance in these samples in which a target adjective modified a projected indicator and tested the agreement of the target's sense with that which the noun was projected to indicate. This procedure tests the sense specificity of the projected indicators in the 100-sentence samples.

The results of this test appear in Table 1 under the heading *Indicator Nouns*. The indicators do turn out to discriminate as projected between target adjective senses, and they do so with 100% reliability. Given this result, the set of indicator nouns can be treated as the basis for a disambiguation procedure. The extent of applicability of such a procedure can be inferred from the Coverage column, which records the proportion of target adjectives that modify projected indicator nouns. Overall coverage is 26.6% (see Table 1), rather low for a disambiguation procedure. The amount of co-occurrence data available for inference has had a substantial effect on coverage. Five of the ten senses are represented by fewer than 100 co-occurrence sentences each, and only one of these five yields any coverage at all. The other five senses are represented by more than 100 sentences each, and every one provides some coverage of the 100-sentence samples. There is therefore every reason to believe that coverage would increase with a larger base for inference. While we found a good semantic matching of adjective senses with the indicators that were recovered from the co-occurrence sentences, the indicator selection depended on an arbitrarily selected 5% level of statistical significance. We therefore investigated the dependence of performance on the chosen significance level. This dependence is highlighted most clearly by comparing the performance of

the statistically significant indicators, listed above, with that of the nouns from the subcorpora that are *not* significant as indicators of target sense. For this purpose, we therefore treated *every* noun from the co-occurrence sentences as an indicator of the sense which that noun is projected to favor in the sample sentences. The results are presented under *All Nouns* in Table 1. Coverage increased from 126 to 237 instances. The 111 newly covered instances are incorrectly assigned in only 7 cases; even when *every* noun from the co-occurrence sentences is treated as an indicator, reliability remains high (97.0%).

The rather high reliability of even those nouns that are not statistically significant indicators of adjective sense suggests that in general text as well as in the co-occurrence sentences, most nouns are highly specific to the sense of their modifying adjectives. For example, not a single color word is a statistically significant indicator for the sense of *light*, although *light blue*, *light brown*, *light gray*, and *light green* all clearly use *light* in its 'not dark' sense. This example also illustrates that many of the individual nouns that we are treating as separate, independent cases actually manifest a smaller number of underlying semantic categories, e.g., *color*. Speakers' knowledge of language must somehow encode such cases, with patterns of use of individual nouns in relation to these adjectives emerging on the basis of that knowledge. A natural way to pursue the necessary revision is in terms of semantic attributes of these nouns, rather than in terms of the nouns themselves. We investigate this possibility, introspectively, in Section 4.4. This was already done to some extent when proper names were grouped together into classes.

The modified noun is not always relevant to the process of disambiguation, and even when the noun is relevant, it is not always sufficient. The observed errors illustrate this (underlined nouns are from the projected indicators):

"I have something hard to speak," he remarked.

A spatula is also used for lifting light pieces of food. . .

They [shells (of bullets)] were small and light, but their turnip shape and radial fins made them difficult to conceal. . .

The auctioneer. . . auctioned off everything, obviously from the estate of an old, dying out family, in short order.

Our spirit is so twisted, torn, because of self, out of its right center, God, and rooted in the flesh; the old life is so foul in the sight of God that no patchwork, no mere polishing up, no amount of varnish will do.

The response to such old masters as Michelangelo, Rembrandt and Velasquez was and still is instant wonder and delight.

The two-inch layer of fat that is attached to the inside of the seal's skin is left intact, and finally the whole hide is turned right side out.

What are the sources of these errors? In the first sentence, *something* is not modified by *hard* at a deep syntactic level; it is instead *to speak something* that relates directly to *hard*, the surface modified noun being simply irrelevant (see Section 4.4.2). In the next case, the noun is not intrinsically irrelevant, but it turns out not to be useful; *pieces* is virtually empty semantically and can be modified by the target adjective in either sense (see Section 5.2). The remaining nouns are relevant. *Family*, *life*, and *master* are ambiguous, and once the ambiguity is resolved the sense of the modifying

adjective is reliably indicated; this issue is addressed in Section 5.1. *Shell* and *side* are also relevant to the sense of the adjective, but even when disambiguated themselves, further information about the context of *light shells* and *right side* is required before the sense of the adjective can be resolved. Section 5.2 addresses some of the contextual relations between adjectives and noun senses that sometimes resolve adjective sense when the intrinsic attributes of the noun sense do not.

4.4 Generalizing the Indicator Nouns

The mutual relevance of nouns and adjectives that permits sense disambiguation is concept specific rather than word specific. More than 40 nouns that are identifiable as indicators of adjective senses reflect a much smaller number of conceptual categories that directly relate to these senses.

4.4.1 Indicator Noun Attributes. The feature *human* provides a useful and conceptually well-motivated basis for interpreting *old*, *right*, and *short*. With one exception (*wine*), the projected indicator nouns for the ‘aged’ sense of *old*—*man*, *people*, *woman*, *you*, *he*, *person*, *lady*, and proper names of people—refer to human beings. Expanding to include all nouns from the co-occurrence sentences that substitute *young* more than *new* for *old*, almost all added nouns continue to be for human beings, as well as certain pronouns (*I*, *we*, *you*, *he*, *she* and *me*, *us*, *her*, *him*), to which animals, plants, and body parts are added. In the 100-sentence sample, all of the 10 ‘aged’ instances of *old* that were not covered by the indicator nouns refer to members of these categories, 7 of them to human beings. Similarly, the correctness sense of *right* refers both to decisions and to decision-making entities, the latter primarily human; among 119 different nouns modified by *right* in the co-occurrence sentences, 14 of the nouns modified by ‘not wrong’ instances are +*human*, and all 55 nouns modified by ‘not left’ instances are –*human*. Finally, in the case of *short*, the vertical extent feature characterizing its ‘not tall’ sense is appropriate to relatively freestanding entities that are normally vertical, and humans are the most talked-about instances of such objects generally. There are differences among the different target adjectives in the appropriateness of the feature. In the case of *old*, it is a restricted version of the feature *living thing*; for *right*, of *animate* (people and animals); and for *short*, human beings happen to be a frequent instantiation of a verticality feature that is normally appropriate to woody plants and to relatively large land animals.

The feature *concrete* is also very widely applicable. All indicators of the ‘not soft’ sense of *hard* are +*concrete*, so –*concrete* reliably indicates the ‘not easy’ sense of *hard*. Because *hard* in its ‘not easy’ sense also modifies concrete nouns syntactically, on the surface (though not semantically; see Section 5), +*concrete* does not as reliably indicate the ‘not soft’ sense of *hard*. Similarly, –*concrete* indicates the ‘not wrong’ senses of *right*, the ‘not long’ senses of *short*, and the ‘gentle’ subset of the ‘not heavy’ senses of *light*. It also indicates ‘not new’ senses of *old*, but in this case –*concrete* is simply a special case of –*animate*.

Another widely relevant class of indicators are *body parts*. These indicate the ‘not young’ sense of *old* and the ‘not long’ sense of *short*. A substantial subset of them indicate the directional sense of *right*. This sense is associated with horizontally separated members of (mostly inherently) paired, repeated entities; it is appropriate to horizontally separated, paired body parts—eyes, ears, thumbs, hands, wrists, arms, legs, feet, etc.—which constitute the majority of nouns, by text frequency, that are modified by the ‘not left’ sense of *right*. For noninherently paired entities, more complex phrasings such as *on the right* or *rightmost* are used instead (also *on the left* or *leftmost*). Body parts are less well represented in the co-occurrence sentences for *hard*; those that occur

(*shell, palate, tissue*) are sense specific for ‘not soft,’ being specific cases of +*concrete* and subject to the reservations on +*concrete* indicator nouns noted above.

Other semantic features are more restricted to individual adjective senses. The attribute +*color* disambiguates about half of the ‘not dark’ instances of *light*. Length-related (‘not long’) senses of *short* are indicated by nouns that are +*time period* (*term, period, day, duration, minute, month, night, time, weekend*, in the co-occurrence sentences), but this attribute is subsumable under –*concrete*. *Text/utterance type* (e.g., *story, note, book, manuscript, monolog, phrase, speech, stanza*), though largely subsumable under –*concrete*, often have +*concrete* realizations (as for *book, note, manuscript*).

Some highly specific attributes are clearly relevant; for example, +*military entity* (*cruiser, carrier, gun, arms, armor, dragoon, flak, machine gun, missile*) indicates the ‘not heavy’ sense of *light*, and +*mental* (*answer, decision, reason, theory, argument, assumption, conclusion, conviction, guess, method, notion, opinion, policy, prediction, proposition, question*) indicates ‘not wrong’ senses of *right*. Until many more adjectives have been investigated, we avoid introducing overly specific features whose range, applicability, and definition may be unclear. These are perhaps subsumable under more general attributes. This is the case for +*mental*, a special case of a highly reliable and more general indicator, –*concrete* (see Section 5.2 concerning the more complex case of *military entity* and the relation of *text type*, which we do use, to the *time period* attribute).

In some cases, semantic features such as those discussed above can be used straightforwardly as sense indicators, just as nouns were. For example, a +*human* noun indicates the ‘not tall’ sense of *short*; a –*concrete* noun indicates a ‘not long’ sense of *short*. In other cases, however, the move from specific nouns to semantic features as sense indicators necessitates the formulation of more specific rules for using them. For example, –*concrete* nouns indicate a ‘not heavy’ sense of *light*, except that –*concrete* nouns that are +*color* indicate its ‘not dark’ sense. This requires a formulation using either feature combinations or rule ordering. We can implement this by first attempting to apply a rule +*color* ⇒ ‘not dark’ and then attempting to apply a rule –*concrete* ⇒ ‘not heavy’, to any unresolved cases (Table 2). More complex rules can be expected to be required in other cases.

We do not propose, in general, to extract automatically either semantic generalizations like those discussed above, or the rules that use them. Informed introspective analysis, aided by a perusal of corpora, seems a surer way toward rule-based formulations. Human analysis and understanding are simply richer than the mechanical statistical tools and data sources presently available for arriving at such rules.

4.4.2 Syntactic Structure and the Adjective–Noun Relation. When the nouns are relevant and indicator nouns are readily recovered, Section 4.4.1 shows that coverage can be increased by exploiting specific indicator nouns in order to infer or to extract automatically general semantic attributes of nouns. For some of the indicators, however, generalization properly takes another course, leading not to semantic but to syntactic cues for sense identification. For example, predicate adjective usage indicates the correctness sense of *right*, which is clearly manifested throughout the APHB Corpus.

That’s not quite right.

*Much thought had gone into that costume, and it seemed just **right** for a poor man’s wife.*

We refer to these as cases of a *predicative* indicator feature.

The most prominent example in our data is generalization from the non-anaphoric *it* indicator for the ‘not easy’ sense of *hard*, which is also applicable for the ‘not wrong’

sense of *right*. This indicator is found in statements of the form *It BE ADJ + infinitival clause*, where *BE* is a form of the verb *to be*. Auxiliary verbs, adverbials, or negation occurs optionally in these statements, and the infinitival clause need not follow immediately. *It* is often deleted, and in such cases an anaphoric pronoun may replace the non-anaphoric *it*, resolving to a preceding or following clause. Also, the verb *BE* may be deleted and the entire construction subordinated to a higher verb such as *seems* or *becomes*. We refer to these as cases of an *infinitival indicator* feature.

*It wasn't **hard** to find Marietta Price.*

*Somehow it has never been **hard** for me to believe in Francis' wounds. . .*

*It's sometimes **hard** for a motorist to pass a young fellow standing on the edge of a highway.*

*Since he's doing this for his physical welfare, it wouldn't be **right** of me to let him be bothered.*

A result of this pattern is that almost any verb will look like an indicator for these same senses of *hard* and *right*.⁶ This appearance, however, is spurious. The proper generalization is simply the syntactic construction, including its variants in which the non-anaphoric *it* does not occur. The variants include those in which the entire infinitival clause, or a gerundive phrase based on it, serves as the subject of the main clause, with *hard* or *right* as predicate adjective,

*But becoming more independent is **hard** for many children.*

and those in which a noun object from the infinitival clause is promoted to serve as subject of the verb of being, in place of *it*.

*Later, Mama may have regretted being married, because Papa was so **hard** to understand.*

*Her energy was tremendous, her scruples **hard** to find.*

Since the adjective is characterizing an action or state of affairs, these cases can be subsumed under the +*activity* or -*concrete* semantic attributes discussed in Section 4.4.1 as indicators of these senses of *hard* and *right*. Even as a premodifier of a noun, the adjectives in this construction often relate semantically to the verb phrase, e.g.,

*This is a **hard** program to carry out.*

What is hard is the carrying out of the program, not the program itself. Once again, it is the syntactic construction, and not the modified noun, that is the relevant indicator.

These considerations help us to refine our use of the adjective-noun relation itself and to put it on a firmer linguistic footing. The adjective-noun relation is directly pertinent to semantic attributes of both the adjective and the noun only when there is a *deep* syntactic relation between them. In the case of the infinitival and related constructions, no such relation holds; the noun modified by the adjective at the surface level

⁶ A minority of verbs with a specific relation to the opposite sense do exist, e.g., *feel* relates to the 'not soft' sense of *hard*. Such verbs relate to the alternative sense only when they are outside of the infinitival constructions.

is irrelevant. A principled approach to adjective disambiguation using nouns therefore requires a determination that the adjective modifies the noun at a deep syntactic level. It is therefore important to take into account the infinitival construction prior to disambiguating *any* adjective—even those for which it does not constitute an indicator.

The last two columns of Table 1 present the results of adjective disambiguation by a combination of syntactic and semantic indicator attributes. The disambiguating rules we used are given in Table 2. The syntactic indicator attributes, *predicative* and *infinitival*, were applied first. Afterward, if a target adjective sense was not resolved, semantic indicator attributes were applied; no individual indicator nouns were used. The semantic attributes that were applied were *animate*, *body part*, *color*, *concrete*, *human*, and *text type*; Church and Hanks (1989) had pointed to two of these attributes, *person* and *body part* (also *time*, previously mentioned above) in a seemingly casual listing of just five attributes potentially useful for describing the lexico-syntactic regularities of noun-verb relations.

Table 1 shows that these few, general attributes cover almost three-quarters of all instances of the target adjectives. Disambiguation by these syntactic and semantic attributes is effectively as reliable as disambiguation using significant indicator nouns: having three apparent errors in disambiguation is not significantly worse than the errorless performance of the significant indicator nouns in the 100-sentence samples. In fact, under a deeper analysis, these three cases are consistent with the pertinent attributes and should not be treated as errors at all. In one sentence,

In contrast to his rangy sons, he was a short, heavy, oaken-barrel sort of man.

short modifies *sort* (–*concrete*) and was thus assigned the sense ‘not long.’ However, it is actually relevant not to the head of the noun phrase, *sort*, but rather to *man* (+*animate*); so treated, *short* would be correctly assigned to ‘not tall.’ More complex are the two instances of *old wine*. In the APHB Corpus at large, the *new/old* contrast applied to *wine* relates chiefly to contexts of production of the wine or of the introduction of a type of wine. The *young/old* contrast relates instead to the maturation of some wines, or more generally, to the developmental phases through which wine passes while aging over a period of years. It is a cultural (and thus semantic) fact that wines and other nonanimate entities that undergo developmental changes and pass through maturational stages are treated as living things. Since the two instances of *wine* in the 100-sentence samples are of this sort, their *old* modifiers are properly assigned to ‘not young’; we assigned them to ‘not new’ under a literal interpretation of –*animate*.

The rules in Table 2 can be easily implemented. The approach presupposes that the natural language processing system within which it is applied includes a reliable, wide-coverage parser to determine the noun phrase modified by an adjective and the head of that noun phrase. The lexical database used by this parser must include semantic attribute tags. Most of those used in this paper are already present in some available machine-readable dictionaries, such as the *Longman Dictionary of Contemporary English*. Such a disambiguation procedure is capable of disambiguating, with very high reliability, about three-quarters of the 100-sentence sample instances of the target adjectives we have investigated.

5. Further Considerations

Although the sense clues discussed so far can be readily implemented as a disambiguation procedure, about a quarter of all instances of the adjectives under study were not covered by the rules presented in Table 2. This section addresses these un-

Table 2

Disambiguation of adjectives by syntactic and semantic attributes. Rules associated with lower numbers are applied before rules associated with higher numbers. Rules associated with the same number are unordered relative to one another.

Target Adjective	Disambiguating Rules
Hard	<ol style="list-style-type: none"> 1. +<i>infinitival</i> ⇒ 'not easy' 2. -<i>concrete</i> ⇒ 'not easy' <li style="padding-left: 2em;">+<i>concrete</i> ⇒ 'not soft'
Light	<ol style="list-style-type: none"> 1. +<i>color</i> ⇒ 'not dark' 2. -<i>concrete</i> ⇒ 'not heavy'
Old	<ol style="list-style-type: none"> 1. -<i>animate</i> ⇒ 'not new' <li style="padding-left: 2em;">+<i>animate</i> ⇒ 'not young'
Right	<ol style="list-style-type: none"> 1. +<i>predicative</i> ⇒ 'not wrong' <li style="padding-left: 2em;">+<i>infinitival</i> ⇒ 'not wrong' 2. -<i>concrete</i> ⇒ 'not wrong' <li style="padding-left: 2em;">+<i>human</i> ⇒ 'not wrong' <li style="padding-left: 2em;">+<i>body part</i> ⇒ 'not left'
Short	<ol style="list-style-type: none"> 1. +<i>human</i> ⇒ 'not tall' <li style="padding-left: 2em;">-<i>concrete</i> ⇒ 'not long' <li style="padding-left: 2em;">+<i>body part</i> ⇒ 'not long' <li style="padding-left: 2em;">+<i>text type</i> ⇒ 'not long'

covered cases. Some are readily characterized in terms of the general approach of the previous section; others are more complex.

The following discussion treats the kinds of properties that systematically relate adjective senses to other features of the sentences in which they occur. Unlike the previous section, it does *not* point to any automated procedure to take advantage of these properties, or to the role they might play in some more encompassing procedure, and it uses coverage and reliability as measures of the actual association of adjective senses with other constructs, irrespective of their recoverability from raw text. Thus, this section is concerned with the nature of underlying relations—not with formulating a disambiguation procedure.

5.1 Indicator Noun Sense Attributes

We have found that a substantial proportion of adjectives can be disambiguated by the nouns they modify, largely on the basis of general semantic attributes characterizing those nouns. These attributes, being semantic, must relate in fact to noun senses and not to nouns per se. This issue is finessed, to some extent, in the projected indicator nouns and thus in our application of attributes based on them. Some attributes happen to apply to all senses of a given noun. For example, in the 100-sentence samples, *course* disambiguates *short*, though once it is used for 'path' and once for 'class,' because both senses are -*concrete*. Some indicator nouns were extracted, not because the attribute applies to all senses of these nouns, but because these nouns are used far more often in senses to which an indicator attribute applies than in those to which it does not apply. For example, *people* shows a statistically significant tendency to be associated with the 'aged' ('not young') senses of *old* (when *people* is the plural of *person*), as judged from the co-occurrence sentences, although one instance of the 'of long standing' ('not new') senses of *old* (when *people* meant 'ethnic group') was also found.

In many instances in the 100-sentence samples, the noun modified by a target adjective was ambiguous with respect to one of the indicator attributes: an indicator attribute did characterize some of the noun's common senses, but not others. For example, the noun *side* is projected to occur equally often with each sense of *right* ('not left' 49.9%, 'not wrong' 50.1%). However, this noun has two broad classes of meanings: one refers to commitments on issues and is *-concrete*; the other refers to flanks and is *+concrete*. As expected on the basis of this semantic attribute, *right* usually means 'not left' when it modifies *side* in its (locational) sense 'flank,' but virtually always means 'not wrong' when modifying *side* in its commitment sense. The sense of *side* is therefore a more reliable indicator of the sense of *right* than is the noun itself.

If these nouns are disambiguated with respect to the relevant attribute, reliability can be increased, as in the case of *right side*. Coverage will increase as well. Some nouns have two or more common senses that disagree in the value of relevant attributes and thus were not recovered as indicator nouns; their *senses* might well be reliable indicator features. Disagreements in the value of a semantic attribute for a given noun can even be systematic. Any noun *N* with sense *S* can be used to mean 'a type of *S*,' as with *family doctor* in

Swedes lament the almost total disappearance of the old family doctor.

When types are construed as *-concrete*, as when referring to roles, such uses are specific to 'not new' senses of *old* (and to 'not wrong' senses of *right*). Thus, any noun or semantic attribute that is associated with the alternative senses of these adjectives would be wrongly interpreted when a 'type of *S*' usage of the noun is not recognized.

Another widely pertinent example is the more complex ambiguity in the reference of modified nouns for roles or relationships. A role noun almost always refers to an individual (*+animate*) who stands in that relationship to another, as in all the example sentences cited below; for example, when *the nuns, new and old, filed out of the cloister*, it was a set of persons and not of relationships who did so. The adjective, however, may apply to that individual (a *+animate* noun sense) or to the role itself (a *-animate* noun sense). Thus, the adjective *old* may apply either to the relationship or to the role designated by *doctor, friend, empress, and nun*, with *old* having the sense 'former' or 'of long standing':

... I was with old friends; I had made new friends; and that night I think that I was lonelier than ever before.

It was only to be expected that the lords and ladies of the court would compare the first wife and the second, the old empress and the new. . . all in favor of the old.

A prayer, the Bishop's blessing—and the nuns, new and old, filed out to the cloister.

or to a person having that relationship or role, with *old* having the sense 'aged':

... he rang to six friends, not too young, not too old, and explained that he'd have to postpone their dinner.

The old doctor and the young doctor rode in silence for two miles and indulged in their memories.

On the contrary, these nuns, young and old, were invariably cheerful and happy, almost gay and full of childish fun and laughter. . .

There is an inherent ambiguity concerning the relation of the adjective *old* to its noun: the referent of the noun is an individual (which is +*animate*), but it is an associated noun sense to which the adjective applies (which may be +*animate* or –*animate*) that strictly determines the sense of *old*. In the 100-sentence samples, when *old* modifies a role noun, it always applies in its ‘aged’ sense to the individual and in its ‘former’ sense to the role. Some role/relationship nouns are used overwhelmingly in their role senses (as with *friend*) or in their personal senses (as with *doctor*). Otherwise, inferring the correct sense for *old* involves a resolution of the function of the noun. For example, an *old forest* is ‘not new’ if it has existed for a great period of time and ‘not young’ if it is in an advanced stage of development in the life cycle of forests (cf., the discussion of *wine* in Section 4.4.2).

This semantic ambiguity in the noun sense to which the adjective applies can therefore be resolved by the same rules formulated for unambiguous cases, once the relevant noun sense is identified. How to access the relevant noun sense is an unsolved problem: the noun’s direct referent is an individual, whereas the semantic structure entailed by the noun is a semantic network, and the adjective may apply to the *network’s* noun sense nodes rather than to the noun referent itself. The utility and elegance of such semantic representations is suggested by linguistic discussions on lexical semantics. They have been used with notable success by Fillmore and Atkins (1991), who exploit the intricacies of such networks in a now-classic account of the semantics of *risk*, with different nodes of the network providing the locus of what might be distilled as the word’s distinct senses. Similar network representations are adopted and implemented in restricted domains in several computational models (e.g., Sowa 1986).

The semantics of noun senses therefore relate more specifically and directly to adjective senses than do nouns themselves; in fact, 38 (30%) of the 125 cases not covered by the rules of Table 2 are resolved when these broader semantic structures are used. In most of these cases, noun senses themselves supplied the attributes used by the rules of Table 2 to disambiguate adjectives. In other cases, such as role nouns, with more complex semantic structures, we are able to resolve the semantic relation of adjective and noun, but this ability cannot be captured in rules as simple as those of Table 2.

5.2 Other Indicators of Adjective Senses

Section 4 showed that indicator nouns and, in particular, certain of their semantic features are quite reliable as bases for interpreting the meanings of the adjectives that modify them. In some cases, however, nouns provided very little assistance—when the pertinent semantic and syntactic features do not apply, the same noun is often simply consistent with alternative senses. This was systematically true for relationship nouns modified by *old*. Similarly, the adjective *light* can refer either to weight or to color in modifying most concrete nouns.

In these cases, disambiguation involves words other than the noun that the target adjective modifies, standing in other syntactic relations to the target. The effectiveness of one such alternative has already been demonstrated—the special case of antonymic adjectives. In the special constructions discussed in Section 3 and, in particular, when they modify the same noun, they disambiguate one another with almost perfect reliability. Consider, for example, the sentence

A piece will seldom bake uniformly, even with the most loving attention—that is, it will vary from light to very dark. . .

The target adjective *light* modifies the pronoun *it*, which refers anaphorically back to

piece (of food being baked). Light pieces of food may be either ‘not dark’ or ‘not heavy,’ so the noun provides no substantial aid to interpretation. However, the phrase *from light to dark* secures the intended sense of *light*.

Verb senses can relate systematically to adjective senses, because adjectives often designate attributes pertinent to the application of the verbal action to/by the referent of the modified object/subject noun. Verbs are therefore useful to the interpretation of adjectives modifying subject or object nouns. Returning to the difficult case of *light*, one of the 7 sentences in which an error was made when all nouns were treated as indicators (see Section 4.3) is

A spatula is also used for lifting light pieces of food. . .

In this sentence, the head noun *pieces* is irrelevant, essentially empty semantically. The critical noun is *food*, but it is not a directly usable indicator; light food may be either ‘not heavy,’ as in this sentence, or ‘not dark,’ as in the previous example. It is instead the verb *lifting* that provides the best sentence-internal indication of the ‘weight’ sense of *light* in the example under consideration. The 5 APHB sentences that refer to the lifting of a light object all involve the ‘not heavy’ sense of *light*. This is not a logical requirement—both dark and heavy objects are also said to be lifted. But the weight of an object is intrinsically relevant and its color irrelevant to the lifting of that object; a reference to the lifting of a *light N* in the ‘not dark’ sense is likely to be misconstrued unless additional cues to interpretation are provided. Similar results are found for semantically similar verbs. Thus, *carry* disambiguates *light* in

Furniture movers, for example, carry light objects in their hands.

whereas the modified noun is of no help. Accordingly, *physically supports* is a semantic attribute of some indicator verbs for the ‘not heavy’ sense of *light*.

In some cases when a noun or even a noun sense is consistent with more than one sense of the target adjective that modifies it, a default target sense may be reliably inferred so long as there is no strong counter-evidence in the immediate context. *Old doctor*, for example, means ‘aged doctor’ in 28 of 30 instances in our corpus. The other 2 cases are both for *old family doctor*. In one, the sentence itself makes it clear that a generic ‘type of doctor’ sense was intended (see Section 5.1). Similarly, *old* means ‘former’ in the sentence

I know that the old family doctor, Dr. Schlomm, always told Manya she could be stabilized on medication, that she could be kept under control.

as shown by the immediately following sentence:

So did her present doctor.

A more complex example of default inference is provided by *right* as a modifier of *side*. The commitment sense of *side* strongly favors the correctness sense of *right*, whereas the locational senses of *side* favor the directional senses of *right*. However, even with the locational sense of *side*, ‘not wrong’ is not an anomalous usage. Compare, for example, the following sentences from the APHB Corpus:

They were hiding behind the big oak on the left side of the road.

*"Can't you see you're on the **wrong side** of the road?"*

*Hoover leaped from his car and ran to the **left side** of the gangster's car.*

*Zaza still stood in the road, on the **wrong side** of the car.*

In the face of such examples, it becomes difficult to interpret the adjective in such sentences as

*But the car, now on the **right** [not left] **side** of the road, was too late to veer away from the second tire. . .*

*He was on the **right** [not wrong] **side** of the screen, he had an excellent day's work behind him, and in two minutes' time he would hear Bing Crosby sing.*

except with respect to the broader discourse context. In our data, locational senses of *side* always involve the directional sense of *right* in *right side of* unless there is decisive evidence to the contrary in the same sentence or in the immediately surrounding discourse. In many of the directional sense uses, we find no overt clue within the sentence or the immediate discourse to determine the sense; it appears to be simply the assumed interpretation when no specific information contradicts it. In contrast, for each of the (6 out of 45) instances in which a correctness sense of *right* modifies a locational sense of *side* in *right side of*, there is something explicit in the near context, usually in the same sentence. For example, the ambiguous case of *the right side of the screen*, above, is resolved by a preceding sentence (two short sentences intervene):

Across the centre of the hall hung a screen, and on this screen was being projected a motion picture; half the men had to see the picture back to front, because they had to look at it from the back [the wrong side] of the screen, but nobody minded that very much.

The directional sense of *right* cannot be construed as a general default, since the correctness sense is far more common overall. Evidently, the locational sense of *side of* is so powerful a semantic indicator for the directional sense of *right* that people do not use *right side of* in other senses without providing substantial countervailing evidence for the sense. Relative to a given indicator noun, then, it makes sense to think of adjective disambiguation in terms of default interpretations.

Some noun-based disambiguation of adjectives involves the noun's functionality rather than its intrinsic semantic attributes; many such nouns relate to relevant attributes of indicator verbs. The adjective *light*, when modifying a +*concrete* noun, is a case in point. Both weight and color characterize all +*concrete* entities. In fact, however, we can reliably determine the sense of *light* in a number of these cases. Three nouns in the co-occurrence sentences emerged as significant indicators for the 'not heavy' sense of *light*—*cruiser*, *load*, and *harness*. It is the weight and not the color of a load that is functionally relevant; the thickness and thus weight of a harness that bears on the speed and load-bearing potential of the draught animals fitted with it; and the heaviness of a vessel that is relevant to the speed, ease of handling, load-bearing potential, and imperviousness to damage that is pertinent to military cruisers. In fact, if it were necessary to specify such entities as being light in color, we would expect that their functional specificity for lightness in weight would lead to the use of a more specific qualifier, such as *light-colored*, rather than simply *light*. Semantic attributes such as *carried things* would characterize nouns such as *load*, e.g., *burden*, *cargo*, or *freight*, and *load-bearing equipment* or *load-relevant equipment* would characterize nouns such as

harness and *cruiser*. In the sample of sentences containing *light*, the following might be subsumed under such attributes: *aircraft*, *brigade*, *car*, *cart*, *defense*, *guard*, *horse*, *industry*, *package*, *shell*, *tank*, and *weight*. Some of these, such as *shell* and *tank*, have meanings to which the load-bearing issue is not relevant, though in the case of *tank* the application of the adjective *light* does appear to restrict its referent to the military vehicle. Others, such as *car*, involve issues of fashion and decoration to which color and thus darkness is potentially relevant as well, although the 5 instances of *light car(s)* in our corpus do refer to weight (with *dark car(s)* referring to a darkened interior). But suitably constrained either to subsets in which decoration is not a functional value (e.g., *military/industrial equipment*) or by treating feature combinations that include, e.g., *-decorative relevance*, it would be possible if not ideal to handle such nouns in terms of attribute values. Thus, some nouns can be disambiguated by relatively narrowly defined semantic classes, such as *military/industrial equipment*, and with high reliability, leading to something close to lexicalized noun phrases, e.g., *light cruiser* or *light industry*. But it is in fact the functional relevance of heaviness versus darkness in the context of its use that is actually involved.

A more complex example is provided by those indicators for the 'not long' sense of *short* that are types of texts or utterances—*book*, *manuscript*, *monologue*, *note*, *phrase*, *poem*, *speech*, *stanza*, *story*, and *syllable*. These, like *+time period*, are largely subsumable under the *-concrete* feature, but there are also *+concrete* instances of some of these nouns, such as *book*, that still select the 'not long' sense of *short*. In these instances, it is not a physical dimension of the item that is short, and reference to such dimensions in the case of *book* relates instead to the 'not tall' sense of *short*. Superficially, this might support the pertinence of a special attribute, *textual*, for disambiguating *short*. However, the 'not long' characterization of shortness of texts refers explicitly or implicitly to the duration of the performance (e.g., reading or reciting) of the text. Accordingly, it is the *time period* attribute that would appear to be involved in this case—an attribute of activities that constitute the typical use of texts, not an attribute of texts themselves. The relationship involved in this case is comparable to that discussed by Pustejovsky and Boguraev (1993), in which a single sense of *fast* relates speed of vehicle motion in both *fast car* and *fast highway* via the qualia structure of lexical entries for *highway* and *car*. Accordingly, although most instances of nouns for text types can disambiguate *short* by being *-concrete*, the principled basis for disambiguating the adjective entails a more complex type of inference than simple characterization of semantic attributes of the modified noun itself.

6. Concluding Remarks

For the adjectives and adjective senses under study, it has been demonstrated that the noun modified by an adjective provides, in most cases, an extremely reliable indication of the sense of that adjective. General semantic attributes of the modified noun provide equally reliable and more widely applicable indications of adjective meanings. These attributes are conceptually relevant and compact ways of representing the semantic relation between adjectives and their modified nouns. When the noun is such that it does not disambiguate the adjective, other words in other specific relations often provide this information. Once again, these words reflect semantic classes rather than specific lexical items and provide a compact and meaningful semantic characterization of the relation of the adjective to the related word class.

This work also supports the view that a small number of close syntactic relations channel much of the semantic interpretation involved in disambiguation, at least in the case of sense dichotomies in adjectives. We suspect that when this highly struc-

tured type of evidence is removed, the more diffuse kinds of clues implicit in mass contextual comparison approaches still have a principled role in disambiguation, e.g., by providing a generalized feel for the overall topical preferences of target senses.

The data on which this study is based were analyzed statistically, and our results were summarized quantitatively. These analyses, however, were not provided in anticipation of the creation of statistical procedures for automated word sense disambiguation, but rather to help us to explore and to establish a property of language—that general semantic attributes of noun senses correlate with the senses of the adjectives that modify these nouns. The phenomenon itself can now be formulated in nonquantitative terms. In addressing or utilizing this phenomenon, corpora remain quite useful, supplying examples of word use that can be used for insight and serious discourse analysis.

The main goals of this paper were to investigate conceptual issues in adjective disambiguation. The above results, however, may provide a basis for a useful core component of an automated procedure that disambiguates adjectives, with the rest to be covered by an auxiliary procedure. This core component's coverage of three-quarters of adjective instances is lower than can be attained by other approaches, even those with quite high reliability. Nonetheless, such a core has practical advantages: when it is applicable, it is virtually errorless; and if there are errors, it is known where they will be concentrated—in the remaining quarter of cases, those resolved by the auxiliary procedure. The full procedure can then attain a 97–98% reliability level even if the auxiliary procedure has only 90% reliability.

The results of this study need to be extended in several ways. First, they should be extended to cover a substantial number of adjectives. Any such enterprise always leads to improved understanding, but there is also reason to believe that the results will be comparable to those we report. The five adjectives investigated in this paper represent a fair range in terms of difficulty: the relations between the pairs of senses range from barely distinct (*short*), to interrelated in complex and sometimes subtle ways (*old*), to semantically orthogonal (*light*). These adjectives are among the most frequent in English and, correspondingly, are applicable to a variety of nouns, so they may be on average somewhat more useful than most for investigating the interpretation of adjectives based upon nouns. It is of course important to investigate types of sense discriminations other than the antonymically defined pairs that were imposed by the design of this study.

The results should also be extended to investigate more fully the contributions of other types of syntactic relations of adjectives. Although it appears that verb-based disambiguation of adjectives can be formulated in the same way as noun-based disambiguation, we have not yet systematically studied the contribution of verbs. As in the case of nouns, we expect the appropriate representation to be in terms of semantic classes rather than sets of words (cf. Levin 1993). It would also be useful to address the possible contribution of adjectives other than the sense-specific antonyms.

Finally, the results need to be extended beyond the class of adjectives. Most directly, nouns are probably as reliably disambiguated by adjectives as adjectives are by nouns.

Appendix A: Extracting Word Sense Indicator Features

This paper treats the disambiguation of adjectives relative to antonym-specific senses, with the nouns they modify as clues suggesting which sense pertains to a given instance. We model this situation more generally as follows.

A.1 Definition

Let T be a target feature of a sentence or other textual unit that is to be disambiguated with respect to some set of characteristics S_1, \dots, S_n ; in our study, n is 2, and S_1 and S_2 are two groups of senses, each group relating semantically to a particular antonym of a target adjective T . Let F be some feature that can be associated with the target feature T in the same textual unit; F may in general be any feature of a sentence or other textual unit that contains T , e.g., a semantic attribute, a word, a phrase, or a grammatical pattern. In our study, F is a noun modified by a target adjective T . We abbreviate the event "sense of $T = S$ " as " S " and " F is in the context of T " as " F ." Then F provides evidence concerning the sense of T if, for some i and j , the conditional probability $P\{S_i | F\}$ is not equal to $P\{S_j | F\}$. We call F an *indicator* for the sense S_i of T if the conditional probability of S_i given F is higher than the conditional probability of any other sense given F : $P\{S_i | F\} > P\{S_j | F\}$ for all $j \neq i$.

A.2 Estimation of $P\{S_i | F\}$

The desired probabilities $P\{S_i | F\}$, $i = 1, \dots, n$, can be estimated from a representative sample of sentences with T in a corpus. However, a subcorpus in which T is disambiguated by a particular set of clues may bias the sense distribution; the desired probabilities differ, in general, from probabilities $Q\{S_i | F\}$ defined on such disambiguated subcorpora. The formula we derive for $P\{S_i | F\}$ takes into account this possible bias. In the derivation, we use an intuitively reasonable assumption that the occurrence of a potential indicator feature F depends only upon the sense of the target T with which it is associated and not upon the fact that it occurs in the subcorpora (in our case, primarily a phrasal substitution context of antonym co-occurrence sentences), i.e., that $Q\{F | S_i\} = P\{F | S_i\}$. This need not be true for arbitrary disambiguating features, but we treat it as approximately correct in the case of nouns modified by a target adjective.

Starting with $P\{S_i | F\}$ and applying the identities

$$P\{S_i | F\} = \frac{P\{F | S_i\}P\{S_i\}}{P\{F\}}, Q\{F | S_i\} = \frac{Q\{S_i | F\}Q\{F\}}{Q\{S_i\}}, \text{ and } \sum_i P\{S_i | F\} = 1,$$

we obtain for $i = 1, \dots, n$

$$P\{S_i | F\} = \frac{q_i}{\sum_k q_k}, \text{ where } q_i = \rho_i Q\{S_i | F\} \text{ and } \rho_i = \frac{P\{S_i\}}{Q\{S_i\}}.$$

The q_i are the conditional probabilities $Q\{S_i | F\}$ weighted by the bias ratios ρ_i defined above. When there is no bias, i.e., when $\rho_i = 1$ and $P\{S_i\} = Q\{S_i\}$ for all i , then the formula yields simply $P\{S_i | F\} = Q\{S_i | F\}$.

For cases with two senses, the formulas are

$$P\{S_1 | F\} = \frac{\rho_1 Q\{S_1 | F\}}{\rho_1 Q\{S_1 | F\} + \rho_2 Q\{S_2 | F\}} \text{ and } P\{S_2 | F\} = \frac{\rho_2 Q\{S_2 | F\}}{\rho_1 Q\{S_1 | F\} + \rho_2 Q\{S_2 | F\}}.$$

Most of the target adjectives we investigated have a substantial sense bias, because their antonyms co-occur with them at substantially different rates. When all instances of F in the disambiguated subcorpora fall in the same sense group, the estimated probability for that sense given the feature F is 1, regardless of the values of the biases ρ_i ; this was the typical case in our experiment because of the specificity of the modified nouns (Section 4.2). But whether or not these estimated sense probabilities are statistically significant does depend on the ρ_i values.

We illustrate the use of the formulas for our data. Let $T = \textit{old}$, $S_1 = \textit{'not new'}$, $S_2 = \textit{'not young'}$, and $F = \textit{friend(s)}$. Because we want to estimate the probabilities $P\{S_i | F\}$ for our sample sentences in particular and not for the corpus at large, it is for these sentences that we determine the $P\{S_i\}$ needed for computing the values of the ρ_i . The data are given in Table 3. In this table, n_1 is the number of instances of sense S_1 of a target in the 100-sentence sample for that target, and n_2 is the number of such instances for the other sense; $P\{S_i | F\}$ is the corresponding probability of the sense S_i of the target in the samples. For our example,

$$P\{S_1\} = \frac{n_1}{n_1 + n_2} = \frac{48}{48 + 44} = 0.522 \text{ and } P\{S_2\} = 1 - P\{S_1\} = 0.478.$$

Estimates for $Q\{S_1\}$ and $Q\{S_2\}$ are obtained from m_1 and m_2 , the numbers of instances in which the two antonyms modify the same noun as the target in the co-occurrence sentences; they are

$$Q\{S_1\} = \frac{m_1}{m_1 + m_2} = \frac{463}{463 + 266} = 0.635 \text{ and } Q\{S_2\} = 1 - Q\{S_1\} = 0.365.$$

The resulting bias ratios are

$$\rho_1 = \frac{P\{S_1\}}{Q\{S_1\}} = \frac{0.522}{0.635} = 0.822 \text{ and } \rho_2 = \frac{P\{S_2\}}{Q\{S_2\}} = \frac{0.478}{0.365} = 1.311.$$

The bias ratios for all five pairs of adjective senses in our study are given in Table 3. The probabilities $Q\{S_1 | F\}$ and $Q\{S_2 | F\}$ are estimated from the observed frequencies in the co-occurrence sentences. *Friend* or *friends* is modified by both *old* and *new* in 9 sentences and by both *old* and *young* in 1 sentence, yielding an estimate of 9/10 for $Q\{S_1 | F\}$ and 1/10 for $Q\{S_2 | F\}$. Substituting these figures into the formula yields the estimates

$$P\{S_1 | F\} = \frac{0.822 \times 0.9}{0.822 \times 0.9 + 1.311 \times 0.1} = 0.849 \text{ and } P\{S_2 | F\} = 1 - 0.849 = 0.151.$$

According to these probabilities we estimate for the sample sentences, we expect that *friend* should appear to indicate the 'not new' sense of *old* there, though somewhat less than in the co-occurrence sentences themselves.

A.3 Assessing Statistical Significance of Indicators

While the estimated probabilities of the senses of *old* when it modifies *friend(s)* do differ substantially, these estimates are based on only 10 occurrences of *friend(s)* in the co-occurrence sentences. The low number of instances of *old* modifying *friend(s)* raises the possibility that the difference in estimated probabilities might be spurious, a chance deviation. This section explains how we tested this possibility, i.e., how we decided whether an observed difference is statistically significant.

Our null hypothesis is that there is no difference between the probabilities $P\{S_1 | F\}$ and $P\{S_2 | F\}$. We test this equiprobability hypothesis using a 5% significance level, i.e., we reject it if the probability of so large an observed deviation from this hypothesis is 5% or less. Because we did not specify a hypothetical sense preference in advance for each word, we use two-tailed tests.

We frame the null hypothesis of equal sense probabilities $P\{S_i | F\}$ in terms of the corresponding sense distribution in the sample sentences. We are testing, however,

Table 3

Projected probabilities and bias factors for five ambiguous adjectives. S_1 and S_2 are the antonym-related sense groups for the target adjective, in alphabetical order by antonym; the m_i ($i = 1, 2$) are the numbers of instances in which the target in sense S_i and the corresponding antonym co-occur modifying the same noun; the $Q\{S_i\}$ are the relative frequencies, $m_i/(m_1 + m_2)$; the n_i are the numbers of instances of the target adjective in sense S_i in the 100-sentence sample for that target; the $P\{S_i\}$ are the relative frequencies, $n_i/(n_1 + n_2)$; the ρ_i are the bias ratios, $P\{S_i\}/Q\{S_i\}$; and the $\tilde{Q}\{S_i\}$ are the probabilities projected for the co-occurrence sentences under the null hypothesis of equiprobability in the general corpus represented by the 100-sentence samples.

Target Adjective	m_1	m_2	$Q\{S_1\}$	$Q\{S_2\}$	n_1	n_2	$P\{S_1\}$	$P\{S_2\}$	ρ_1	ρ_2	$\tilde{Q}\{S_1\}$	$\tilde{Q}\{S_2\}$
Hard	14	43	0.246	0.754	69	20	0.775	0.225	3.157	0.298	0.086	0.914
Light	59	33	0.641	0.359	31	67	0.316	0.684	0.493	1.906	0.794	0.206
Old	463	266	0.635	0.365	48	44	0.522	0.478	0.822	1.311	0.615	0.385
Right	312	142	0.687	0.313	22	77	0.222	0.778	0.323	2.487	0.885	0.115
Short	185	18	0.911	0.089	86	9	0.916	0.084	0.933	1.068	0.518	0.482

in the co-occurrence sentences. We need to determine the probabilities $Q\{S_i | F\}$ for the co-occurrence sentences that would correspond to $P\{S_i | F\} = 1/2$ in the corpus at large. We denote this probability as $\tilde{Q}\{S_i | F\}$. We use the same formula for this computation as was used in the estimation problem above. This time, however, it is the probability in the co-occurrence sentences that we project from an assumed 50–50 distribution of senses with respect to the noun for the sample sentences; i.e., for cases with two senses, we have to solve the simultaneous equations

$$\frac{1}{2} = \frac{\rho_1 \tilde{Q}\{S_1 | F\}}{\rho_1 \tilde{Q}\{S_1 | F\} + \rho_2 \tilde{Q}\{S_2 | F\}} \quad \text{and} \quad \frac{1}{2} = \frac{\rho_2 \tilde{Q}\{S_2 | F\}}{\rho_1 \tilde{Q}\{S_1 | F\} + \rho_2 \tilde{Q}\{S_2 | F\}}$$

with $\tilde{Q}\{S_1 | F\}$ and $\tilde{Q}\{S_2 | F\}$ as unknowns. The values

$$\tilde{Q}\{S_1 | F\} = \frac{\rho_2}{\rho_1 + \rho_2} \quad \text{and} \quad \tilde{Q}\{S_2 | F\} = \frac{\rho_1}{\rho_1 + \rho_2}$$

are unique solutions to these equations.⁷

Returning to the example of *old friend(s)*, the null hypothesis is reformulated as

$$\tilde{Q}\{S_1 | F\} = \frac{1.311}{0.822 + 1.311} = 0.615 \quad \text{and} \quad \tilde{Q}\{S_2 | F\} = 1 - 0.615 = 0.385.$$

We now have a model for the probability of each sense in the co-occurrence sentences, on the assumption that sense selection is in fact independent of the modified noun

⁷ In the general case, if the target has n senses, the equiprobability hypothesis is expressed by the simultaneous equations

$$\frac{1}{n} = \frac{\rho_i \tilde{Q}\{S_i | F\}}{\sum_k \rho_k \tilde{Q}\{S_k | F\}} \quad i = 1, \dots, n,$$

with the unique solutions

$$\tilde{Q}\{S_i | F\} = \frac{\frac{1}{\rho_i}}{\sum_{k=1}^n \frac{1}{\rho_k}} \quad i = 1, \dots, n.$$

in the corpus at large. In the example of *old friend(s)*, with $\tilde{Q}\{S_1 | F\} = 0.615$, 6.15 of the 10 instances of *old friend(s)* in the co-occurrence sentences are expected to be 'not new.' The 9 attested instances, then, are somewhat greater than expected under the null hypothesis. To decide whether this excess might be a chance deviation from expectation, we compute the probability p that 9 or more instances of 'not young' would be observed out of 10 instances of *old friend(s)*, if $Q\{S_1 | F\} = 0.615$. Since we are using a two-tailed test with a significance level of 5%, we reject the null hypothesis if $p < 0.025$.

The formula for computing the one-tailed probability p is a sum of binomial probabilities. Let S_i be the sense that we project to be preferred in the corpus at large, given that the modified noun is F ; i.e., it is the sense S_1 if $P\{S_1 | F\} > P\{S_2 | F\}$, and it is the sense S_2 otherwise. Then

$$p = \sum_{k=m_t}^{m_1+m_2} \binom{m_1+m_2}{k} \times \tilde{Q}\{S_t | F\}^k \times (1 - \tilde{Q}\{S_t | F\})^{m_1+m_2-k},$$

where the m_i are the numbers of instances of the target in sense S_i that modify F in the co-occurrence sentences. In the example of *old friend(s)*, $m_1+m_2 = 10$, $\tilde{Q}\{S_1 | F\} = 0.615$, $t = 1$, and $m_t = 9$; so the probability of observing 9 or more instances of *old friend(s)* in the sense 'not new,' in the co-occurrence sentences, is $\sum_{k=9}^{10} \binom{10}{k} \times 0.615^k \times 0.385^{10-k} = 0.056$ under the null hypothesis. Since p is greater than 0.025, we do not reject the null hypothesis; *friend(s)* is not a statistically significant indicator for either sense of *old*.

When all instances of F fall in the same sense group (which was typically the case in our data), the formula for one-tailed probability p is reduced to the one-term expression

$$p = \tilde{Q}\{S_t | F\}^{m_t}.$$

In the case of *old person(s)*, $m_1 = 0$, $m_2 = 7$, $P\{S_1 | F\} = 0$, $P\{S_2 | F\} = 1$, $t = 2$, and $\tilde{Q}\{S_2 | F\} = 0.385$; so $p = 0.385^7 = 0.00125$. Since p is less than 0.025, we reject the null hypothesis of equiprobability of senses; *person(s)* is a statistically significant indicator for the 'not young' sense of *old*.

Acknowledgments

We thank Branimir Boguraev, Judith Klavans, and two anonymous reviewers for very helpful comments on earlier versions of this paper.

References

- Boguraev, Branimir; Byrd, Roy; Klavans, Judith; and Neff, Mary (1989). "From structural analysis of lexical resources to semantics in a lexical knowledge base." *Proceedings, Workshop on Lexical Acquisition. International Joint Conference on Artificial Intelligence (IJCAI)*.
- Braden-Harder, Lisa (1991). "Sense disambiguation using on-line dictionaries: An approach using linguistic structure and multiple sources of information." Masters thesis, Department of Computer Science, New York University, New York.
- Charles, Walter G., and Miller, George A. (1989). "Contexts of antonymous adjectives." *Applied Psycholinguistics* 10(3):357-375.
- Church, Kenneth W., and Hanks, Patrick (1989). "Word association norms, mutual information, and lexicography." In *Proceedings, 27th Annual Meeting of the Association for Computational Linguistics*, ACL. 76-83.
- Fillmore, Charles A., and Atkins, Beryl T. (1991). "Word meaning: Starting where MRDs stop." Plenary paper presented at the 29th Annual Meeting of the Association for Computational Linguistics, Berkeley, California.
- Francis, Winthrop N., and Kucera, Henry

- (1982). *Frequency Analysis of English Usage: Lexicon and Grammar*. Boston: Houghton-Mifflin.
- Gale, William; Church, Kenneth W.; and Yarowsky, David (1992). "Estimating upper and lower bounds on the performance of word-sense disambiguation programs." *Proceedings, 30th Annual Meeting of the Association for Computational Linguistics, ACL*. 249-256.
- Gross, Derek; Fischer, Ute; and Miller, George A. (1989). "Antonymy and the representation of adjectival meanings." *Journal of Memory and Language* 28(1):92-106.
- Hindle, Donald, and Rooth, Mats (1991). "Structural ambiguity and lexical relations." *Proceedings, 29th Annual Meeting of the Association for Computational Linguistics, ACL*. 229-236.
- Justeson, John S., and Katz, Slava M. (1991). "Co-occurrences of antonymous adjectives and their contexts." *Computational Linguistics* 17(1):1-19.
- Justeson, John S., and Katz, Slava M. (1992). "Redefining antonymy: The textual structure of a semantic relation." *Journal of Literary and Linguistic Computing* 7(3):176-184.
- Levin, Beth (1993). *English Verb Classes and Alternations: A Preliminary Investigation*. Chicago: University of Chicago Press.
- Maarek, Yoelle, and Smadja, Frank (1989). "Full text indexing based on lexical relations, an application: Software libraries." *Proceedings, ACM SIGIR, ACM*. 198-206.
- Pustejovsky, James, and Boguraev, Branimir (1993). "Lexical knowledge representation and natural language processing." *Artificial Intelligence. Special Issue on Natural Language* 63:193-223.
- Sowa, John F. (1986). *Conceptual Structures: Information Processing in Mind and Machine*. Reading, Massachusetts: Addison-Wesley.
- Yarowsky, David (1992). "Word-sense disambiguation using statistical models of Roget's categories trained on large corpora." In *Proceedings, International Conference on Computational Linguistics (COLING-92)*. 454-460.

