Learning Subjective Adjectives from Corpora

Janyce M. Wiebe
Department of Computer Science
New Mexico State University, Las Cruces, NM 88003
wiebe@cs.nmsu.edu

Abstract

Subjectivity tagging is distinguishing sentences used to present opinions and evaluations from sentences used to objectively present factual information. There are numerous applications for which subjectivity tagging is relevant, including information extraction and information retrieval. This paper identifies strong clues of subjectivity using the results of a method for clustering words according to distributional similarity (Lin 1998), seeded by a small amount of detailed manual annotation. These features are then further refined with the addition of lexical semantic features of adjectives, specifically *polarity* and *gradability* (Hatzivassiloglou & McKeown 1997), which can be automatically learned from corpora. In 10-fold cross validation experiments, features based on both similarity clusters and the lexical semantic features are shown to have higher precision than features based on each alone.

Introduction

Subjectivity in natural language refers to aspects of language used to express opinions and evaluations (Banfield 1982; Wiebe 1994). *Subjectivity tagging* is distinguishing sentences used to present opinions and other forms of subjectivity (*subjective sentences*) from sentences used to objectively present factual information (*objective sentences*). This task is especially relevant for news reporting and Internet forums, in which opinions of various agents are expressed. There are numerous applications for which subjectivity tagging is relevant. Two examples are information extraction and information retrieval. Assigning subjectivity labels to documents or portions of documents is an example of a non-topical characterization of information. Current information extraction and retrieval technology focuses almost exclusively on the subject matter of the documents. However, additional components of a document influence its relevance to particular users or tasks, including, for example, the evidential status of the material presented, and attitudes adopted in favor of or against a particular person, event, or position.[1] A summarization system would benefit from distinguishing sentences intended to present factual material from those intended to present opinions, since many summaries are meant to include only facts. In the realm of Internet forums, subjectivity judgements would be useful for recognizing inflammatory messages ("flames") and mining on-line forums for reviews of products and services.

To use subjectivity tagging in applications, good linguistic clues must be found. As with many pragmatic and discourse distinctions, existing lexical resources such as machine readable dictionaries (Procter 1978) and ontologies for natural language processing (NLP) (Mahesh & Nirenburg 1995; Hovy 1998), while useful, are not sufficient for identifying such linguistic clues, because they are not comprehensively coded for subjectivity. This paper addresses learning subjectivity clues from a corpus.

Previous work on subjectivity (Wiebe, Bruce, & O'Hara 1999; Bruce & Wiebe 2000) established a positive and statistically significant correlation with the presence of adjectives. Since the mere presence of one or more adjectives is useful for predicting that a sentence is subjective, this paper uses the performance of the simple adjective feature as a baseline, and identifies higher quality adjective features using the results of a method for clustering words according to distributional similarity (Lin 1998), seeded by a small amount of detailed manual annotation. These features are then further refined with the addition of lexical semantic features of adjectives, specifically *polarity* and *gradability* (Hatzivassiloglou & McKeown 1997), which can be automatically learned from corpora. In 10-fold cross validation experiments, features based on both similarity clusters and the lexical semantic features are shown to have higher precision than features based on each alone. The new adjective features are available on the Web at http://www.cs.nmsu.edu/˜wiebe.

In the remainder of this paper, subjectivity and previous work on automatic subjectivity tagging are first described. The statistical techniques used to create the new adjective features are described next, starting with distributional similarity, followed by learning gradable and polar adjectives. The results are then presented, followed by the conclusions.

Subjectivity

Sentence (1) is an example of a simple subjective sentence, and (2) is an example of a simple objective sentence: [2]

Copyright © 2000, American Association for Artificial Intelligence (www.aaai.org). All rights reserved.

[1]This point is due to Vasileios Hatzivassiloglou.

[2]Due to space limitations, this section glosses over some important distinctions involving subjectivity. The term *subjectivity* is due to Ann Banfield (1982), though I have changed its

(1)
At several different layers, it's a fascinating tale. *Subjective sentence.*

(2)
Bell Industries Inc. increased its quarterly to 10 cents from 7 cents a share. *Objective sentence.*

Sentences (3) and (4) illustrate the fact that sentences about speech events may be subjective or objective:

(3)
Northwest Airlines settled the remaining lawsuits filed on behalf of 156 people killed in a 1987 crash, but claims against the jetliner's maker are being pursued, a federal judge said. *Objective sentence.*

(4)
"The cost of health care is eroding our standard of living and sapping industrial strength," complains Walter Maher, a Chrysler health-and-benefits specialist. *Subjective sentence.*

In (3), the material about lawsuits and claims is presented as factual information, and a federal judge is given as the source of information. In (4), in contrast, a complaint is presented. An NLP system performing information extraction on (4) should not treat the material in the quoted string as factual information, with the complainer as a source of information, whereas a corresponding treatment of sentence (3) would be fine.

Subjective sentences often contain individual expressions of subjectivity. Examples are *fascinating* in (1), and *eroding*, *sapping*, and *complains* in (4). The following paragraphs mention aspects of subjectivity expressions that are relevant for NLP applications.

First, although some expressions, such as *!*, are subjective in all contexts, many, such as *sapping* and *eroding*, may or may not be subjective, depending on the context in which they appear. A *potential subjective element* is a linguistic element that may be used to express subjectivity. A *subjective element* is an instance of a potential subjective element, in a particular context, that is indeed subjective in that context (Wiebe 1994).

Second, there are different types of subjectivity. This work focuses primarily on three: positive evaluation (e.g., *fascinating*), negative evaluation (e.g., *terrible*), and speculation (e.g., *probably*).

Third, a subjective element expresses the subjectivity of a *source*, who may be the writer or someone mentioned in the text. For example, the source of *fascinating* in (1) is the writer, while the source of the subjective elements in (4) is Maher. In addition, a subjective element has a *target*, i.e., what the subjectivity is about or directed toward. In (1), the target is a tale; in (2), the target is the cost of health care. These are examples of *object-centric subjectivity*, which is about an object mentioned in the text (other examples: "I love this project"; "The software is horrible"). Subjectivity may also be *addressee-oriented*, i.e., directed toward the listener or reader (e.g., "You are an idiot").

There may be multiple subjective elements in a sentence, possibly of different types and attributed to different sources and targets. As described below, individual subjective elements were annotated as part of this work, refining previous work on sentence-level annotation.

With colleagues, I am pursuing three applications in related work: recognizing flames, mining Internet forums for product reviews, and clustering messages by ideological point of view (i.e., clustering messages into "camps"). There has been work on these applications: Spertus (1997) developed a flame-recognition system that relies on a small number of complex clues; Terveen et al. (1997) developed a system that mines news groups for Web page recommendations; Sack (1995) developed a knowledge-based system for recognizing ideological points of view; and Kleinberg (1998) discussed using hyperlink connectivity for this problem. Our approach is meant to supplement such approaches: we are developing a repository of potential subjective elements to enable us to exploit subjective language in these applications. This paper takes a significant step in that direction by demonstrating a process for learning potential subjective elements from corpora.

Previous Work on Subjectivity Tagging

In previous work (Wiebe, Bruce, & O'Hara 1999; Bruce & Wiebe 2000), a corpus of 1,001 sentences[3] of the Wall Street Journal Treebank Corpus (Marcus *et al.* 1993) was manually annotated with subjectivity classifications. Specifically, three humans assigned a subjective or objective label to each sentence. They were instructed to consider a sentence to be subjective if they perceived any significant expression of subjectivity (of any source), and to consider the sentence to be objective, otherwise. The EM learning algorithm was used to produce corrected tags representing the consensus opinions of the judges (Goodman 1974; Dawid & Skene 1979). The total number of subjective sentences in the data is 486 and the total number of objective sentences is 515.

In (Bruce & Wiebe 2000), a statistical analysis of the assigned classifications was performed, showing that adjectives are statistically significantly and positively correlated with subjective sentences in the corpus on the basis of the log-likelihood ratio test statistic G^2. The probability that a sentence is subjective, simply given that there is at least one adjective in the sentence, is 55.8%, even though there are more objective than subjective sentences in the corpus.

An automatic system to perform subjectivity tagging was developed as part of the work reported in (Wiebe, Bruce, & O'Hara 1999). In 10-fold cross validation experiments applied to the corpus described above, a probabilistic classifier obtained an average accuracy on subjectivity tagging of 72.17%, more than 20 percentage points higher than a

meaning somewhat to adapt it to this work. For references to work on subjectivity, please see (Banfield 1982; Fludernik 1993; Wiebe 1994).

[3]Compound sentences were manually segmented into their conjuncts, and each conjunct is treated as a separate sentence.

baseline accuracy obtained by always choosing the more frequent class. Five part-of-speech features, two lexical features, and a paragraph feature were used. An analysis of the system showed that the adjective feature was important for realizing the improvements over the baseline accuracy.

Experiments

In this paper, the corpus described above is used, augmented with new manual annotations. Specifically, given the sentences classified as subjective in (Wiebe, Bruce, & O'Hara 1999), the annotators were asked to identify the subjective elements in the sentence, i.e., the expressions they feel are responsible for the subjective classification. They were also asked to rate the strength of the elements (on a scale of 1 to 3, with 3 being the strongest). The subjective element annotations of one judge were used to seed the distributional similarity process described in the next section.

In the experiments below, the precision of a simple prediction method for subjectivity is measured: a sentence is classified as subjective if at least one member of a set of adjectives S occurs in the sentence, and objective otherwise. Precision is measured by the conditional probability that a sentence is subjective, given that one or more instances of members of S appears. This metric assesses feature quality: if instances of S appear, how likely is the sentence to be subjective?

Improving Adjective Features Using Distributional Similarity

Using his broad-coverage parser (Lin 1994), Lin (1998) extracts dependency triples from text consisting of two words and the grammatical relationship between them: $(w1, relation, w2)$. To measure similarity between two words $w1$ and $w1'$, the relation-word pairs correlated with $w1$ are identified and the relation-word pairs correlated with $w1'$ are identified. A similarity metric is defined in terms of these two sets. Correlation is measured using the mutual information metric. Lin processed a 64-million corpus consisting of news articles, creating a thesaurus entry for each word consisting of the 200 words of the same part-of-speech that are most similar to it.

The intuition behind this type of process is that words correlated with many of the same things in text are more similar. It is intriguing to speculate that this process might discover functional and pragmatic similarities. I hypothesized that, seeded with strong potential subjective elements, Lin's process would find others, not all of which would be strict synonyms of the seeds.

A challenging test was performed: in 10-fold cross validation experiments, 1/10 of the data was used for training and 9/10 of the data was used for testing. Specifically, the corpus was partitioned into 10 random sets. For each training set i, all adjectives were extracted from subjective elements of strength 3, and, for each, the top 20 entries in Lin's thesaurus entry were identified. These are the *seed sets* for fold i (each seed set also includes the original seed). The seed sets for fold i were evaluated on the test set for fold i, i.e., the entire corpus minus the 1/10 of the data from which the seeds were selected.

As mentioned above, the precision of a simple adjective feature is used as a baseline in this work, specifically the conditional probability that a sentence is subjective, given that at least one adjective appears. The average precision across folds of the baseline adjective feature is 55.8%. The average precision resulting from the above process is 61.2%, an increase of 5.4 percentage points. To compare this process with using an existing knowledge source, the process was repeated, but with the seed's synonyms in WordNet (i.e., the seed's *synset*) (Miller 1990) in place of words from Lin's thesaurus entry. The performance is slightly better with WordNet (62.0%), but the coverage is lower. When the process below (which gives the best results) is applied using the WordNet synsets, the resulting frequencies are very low. While I believe WordNet is potentially a valuable resource for identifying potential subjective elements, Lin's thesaurus entries appear better suited to the current process, because they include looser synonyms than those in WordNet.

Some adjectives that have frequent non-subjective uses are introduced by the above process. Thus, some simple filtering was performed using the training set. For all seed sets for which the precision is less than or equal to the precision of the baseline adjective feature in the training set, the entire seed set was removed. Then, individual adjectives were removed if they appeared at least twice and their precision is less than or equal to the baseline precision on the training set. This results in an average precision of 66.3% on the test sets, 7.5 percentage points higher than the baseline average. The filtered sets are the ones used in the process below.

Refinements with Polarity and Gradability

Hatzivassiloglou and McKeown (1997) present a method for automatically recognizing the *semantic orientation* or *polarity* of adjectives, which is the direction the word deviates from the norm for its semantic group. Words that encode a desirable state (e.g., *beautiful*, *unbiased*) have a positive orientation, while words that represent undesirable states have a negative orientation.

Most antonymous adjectives can be contrasted on the basis of orientation (e.g., *beautiful–ugly*); similarly, nearly synonymous terms are often distinguished by different orientations (e.g., *simple–simplistic*). While orientation applies to many adjectives, there are also those that have no orientation, typically as members of groups of complementary, qualitative terms (Lyons 1977) (e.g., *domestic*, *medical*, or *red*). Since orientation is inherently connected with evaluative judgements, it is a good feature for predicting subjectivity.

Hatzivassiloglou and McKeown's method automatically assigns a $+$ or $-$ orientation label to adjectives known to have some semantic orientation. Their method is based on information extracted from conjunctions between adjectives in a large corpus—because orientation constrains the use of the words in specific contexts (e.g., compare *corrupt and brutal* with **corrupt but brutal*), observed conjunctions of adjectives can be exploited to infer whether the conjoined words are of the same or different orientation. Using a shallow parser in a 21 million word corpus of Wall Street Journal articles, they developed and trained a log-linear statistical model that predicts whether any two adjectives have the same orientation with 82% accuracy. Combining constraints among many adjectives, a clustering algorithm separates the adjectives into groups of different orientations, and, finally, adjectives are labeled positive or negative. Some manual annotation is required for this process.

Gradability is the semantic property that enables a word to participate in comparative constructs and to accept modifying expressions that act as intensifiers or diminishers. Gradable adjectives express properties in varying degrees of strength, relative to a norm either explicitly mentioned or implicitly supplied by the modified noun (for example, a *small planet* is usually much larger than a *large house*). This relativism in the interpretation of gradable words also makes them good predictors of subjectivity.

A method for classifying adjectives as gradable or non-gradable is presented in (Hatzivassiloglou & Wiebe 2000). A shallow parser is used to retrieve all adjectives and their modifiers from a large corpus tagged for part-of-speech with Church's PARTS tagger (Church 1988). Hatzivassiloglou compiled by hand a list of 73 adverbs and noun phrases (such as *a little*, *exceedingly*, *somewhat*, and *very*) that are frequently used as grading modifiers. The number of times each adjective appears modified by a term from this list is a first indicator of gradability. Inflected forms of adjectives in most cases indicate gradability. Thus, a morphological analysis system was implemented to detect inflected forms of adjectives, a second indicator of gradability. A log-linear statistical model was developed to derive a final gradability judgement, based on the two indicators above.

The work reported in this paper uses samples of adjectives identified as having positive polarity, having negative polarity, and being gradable. For each type, we have samples of those assigned manually and samples of those assigned automatically. These samples were determined using a corpus from the Wall Street Journal, but a different corpus from the one used in the current paper. It is important to note that these adjective sets are only samples. Others currently exist, and more could be produced by running the automatic processes on new data.

Results and Discussion

Table 1 presents the results for the seed sets intersected with the gradability and polarity sets (i.e., the lexical sets). Detailed information is given for the features involving the automatically generated lexical sets. Summary information is given in the bottom of the table for the manually classified lexical sets.

The test data for each fold is the entire data set minus the data from which the seeds for that fold were selected. The columns in the table give, in order from the left, the fold number, the number of subjective sentences in the test set (*# Subj*), the number of objective sentences (*# Obj*), and the precision of the baseline adjective feature, i.e., $p(\text{subjective sentence} \mid \text{an adjective})$ (*Adj*). The *Seed-Freq* columns give the number of sentences in the test set that have at least one member of a seed set for that fold and the *Seed-Prec* columns give:
$p(\text{subjective sentence} \mid \text{an adjective in the seed set})$. The *Lex-Freq* columns give the number of sentences that have at least one member of the indicated lexical set, e.g., *Pol+,auto*, and the *Lex-Prec* columns give:
$p(\text{subjective sentence} \mid \text{an adjective in that set})$. For the $S \cap L$ columns, the set is the intersection of the seed sets for that fold and the lexical set. The $S \cap$ *L-Freq* and $S \cap$ *L-Prec* columns are as above. The *Ave diff* lines give the average difference across folds between the precisions of the indicated sets.

For example, in the test set for Fold 1:

***# Subj*:** there are 428 subjective sentences.

***# Obj*:** there are 475 objective sentences.

***Adj*:** the probability that a sentence is subjective, given an adjective, is 55%.

***Seed-Freq*:** 192 sentences contain a member of the seed set.

***Seed-Prec*:** the probability that a sentence is subjective, given an adjective in the seed set, is 58%.

***Lex-Freq*:** 176 sentences contain an automatically identified positive polarity adjective.

***Lex-Prec*:** the probability that a sentence is subjective, given such an adjective, is 59%.

$S \cap$ *L-Freq*: 56 sentences contain an adjective that is both in the seed set and in the set of automatically identified positive polarity adjectives.

$S \cap$ *L-Prec*: the probability that a sentence is subjective, given such an adjective, is 71%.

The results are very promising. In all cases, the average improvement over the baseline of the $S \cap L$ features is at least 9 percentage points. On average, the gradability/polarity sets and the seed sets are more precise together than they are alone (this information can be found in the *Ave diff* lines). There are many excellent individual results, especially among the *Grad,auto* and *Pol-,auto* sets intersected with the seed sets. Only weak filtering of the original seed sets was done using only the original training data. There are more promising ways in which the various sets could be filtered. For example, some of the data that is currently part of the test set could be used to filter the sets (perhaps 3/10 of the data might be used for training with 1/3 of the training data used for seeding, and 2/3 used for filtering).

Conclusions and Future Work

Learning linguistic knowledge from corpora is currently an active and productive area of NLP (e.g., (Lin 1998; Lee 1999; Rooth *et al.* 1999)). These techniques are often used to learn knowledge for semantic tasks. This paper presents a case study of using such techniques to learn knowledge useful for a pragmatic task, subjectivity tagging.

The results of a clustering method (Lin 1998), seeded by a small amount of detailed manual annotation, were used to develop promising adjective features. These features were then further refined with the addition of lexical semantic

				Pol+,auto						Pol−,auto					
				Seed		Lex		S ∩ L		Seed		Lex		S ∩ L	
Fold	# Subj	# Obj	Adj	Freq	Prec	Freq	Prec	Freq	Prec	Freq	Prec	Freq	Prec	Freq	Prec
1	428	475	55	192	58	176	59	56	71	192	58	75	73	18	61
2	433	469	55	148	64	181	60	53	70	148	64	78	73	14	64
3	444	456	56	86	62	180	60	34	65	86	62	78	74	5	80
4	439	463	56	57	70	178	62	29	69	57	70	77	75	70	71
5	436	465	56	166	63	181	60	52	65	166	63	72	76	10	90
6	443	458	56	133	57	183	60	65	62	133	57	75	75	17	65
7	437	464	56	128	66	181	60	47	70	128	66	76	76	12	75
8	442	459	56	226	60	178	61	58	64	226	60	66	73	18	83
9	439	463	56	147	63	183	61	42	62	147	63	73	73	3	67
10	443	463	56	106	70	179	61	40	68	106	70	68	75	9	89
AVE:			55.8	139	63.3	180	60.4	47.6	66.6	139	63.3	73.8	74.3	11.3	74.5
				Ave diff from Adj to (S ∩ L): 10.8 Ave diff from Seed to (S ∩ L): 3.3 Ave diff from Lex to (S ∩ L): 6.2						Ave diff from Adj to (S ∩ L): 18.7 Ave diff from Seed to (S ∩ L): 11.2 Ave diff from Lex to (S ∩ L): 0.2					

				Pol−+,auto						Grad,auto					
				Seed		Lex		S ∩ L		Seed		Lex		S ∩ L	
Fold	# Subj	# Obj	Adj	Freq	Prec	Freq	Prec	Freq	Prec	Freq	Prec	Freq	Prec	Freq	Prec
1	428	475	55	192	58	235	63	73	68	192	58	37	68	8	75
2	433	469	55	148	64	243	64	67	69	148	64	37	68	18	78
3	444	456	56	86	62	242	64	39	67	86	62	46	65	10	70
4	439	463	56	57	70	238	66	36	69	57	70	43	67	11	82
5	436	465	56	166	63	238	64	61	69	166	63	43	70	20	85
6	443	458	56	133	57	242	64	79	62	133	57	41	68	12	83
7	437	464	56	128	66	241	65	59	71	128	66	41	68	18	83
8	442	459	56	226	60	233	64	74	68	226	60	39	72	11	82
9	439	463	56	147	63	241	64	45	62	147	63	39	64	23	70
10	443	463	56	106	70	232	65	49	71	106	70	36	72	16	88
AVE:			55.8	139	63.3	238.5	64.3	58.2	67.6	139	63.3	39.6	68.2	14.7	79.6
				Ave diff from Adj to (S ∩ L): 11.8 Ave diff from Seed to (S ∩ L): 4.3 Ave diff from Lex to (S ∩ L): 3.3						Ave diff from Adj to (S ∩ L): 23.8 Ave diff from Seed to (S ∩ L): 16.3 Ave diff from Lex to (S ∩ L): 11.4					

Pol+,man: Ave diff from Adj to (S ∩ L): 09.1 Ave diff from S to S ∩ L: 01.6 Ave diff from L to (S ∩ L):03.6
Pol−,man: Ave diff from Adj to (S ∩ L): 20.3 Ave diff from S to S ∩ L: 12.8 Ave diff from L to (S ∩ L): 09.1
Grad,man: Ave diff from Adj to (S ∩ L): 13.1 Ave diff from S to S ∩ L: 05.6 Ave diff from L to (S ∩ L): 06.3

Key:
Pol+: positive polarity. **Pol−**: negative polarity. **Grad**: gradable.
Man: manually identified. **Auto**: automatically identified.
 # Subj: number of subjective sentences. **# Obj**: number of objective sentences.
Adj: precision of adjective feature. **Seed**: seed sets. **Lex** : the lexical feature set, e.g., Pol+,auto.
S ∩ L: seed sets ∩ the lexical feature set.

Table 1: Subjectivity Predictability Results

features of adjectives, specifically *polarity* and *gradability* (Hatzivassiloglou & McKeown 1997), which can be automatically learned from corpora. The results are very promising, showing that both processes have the potential to improve the features derived by the other one.

The adjectives learned here are currently being incorporated into a system for recognizing flames in Internet forums. In addition, we plan to apply the methods to a corpus of Internet forums, to customize knowledge acquisition to that genre. This will include deriving a new thesaurus based on distributional similarity, reapplying the processes for identifying gradable and polar adjectives, and annotating subjective elements in the new genre, from which seeds can be selected.

Acknowledgements

This research was supported in part by the Office of Naval Research under grant number N00014-95-1-0776. I am grateful to Vasileios Hatzivassiloglou and Kathy McKeown for sharing the results of their ACL-97 paper; to Dekang Lin for making the results of his COLING-ACL-98 paper available on the Web; to Aravind Joshi for suggesting annotation of subjective elements; to Matthew Bell for performing the subjective element annotations; and to Dan Tappan for performing supporting programming.

References

Banfield, A. 1982. *Unspeakable Sentences*. Boston: Routledge and Kegan Paul.

Bruce, R., and Wiebe, J. 2000. Recognizing subjectivity: A case study of manual tagging. *Natural Language Engineering*.

Church, K. 1988. A stochastic parts porgram and noun phrase parser for unrestricted text. In *Proc. ANLP-88*, 136–143.

Dawid, A. P., and Skene, A. M. 1979. Maximum likelihood estimation of observer error-rates using the EM algorithm. *Applied Statistics* 28:20–28.

Fludernik, M. 1993. *The Fictions of Language and the Languages of Fiction*. London: Routledge.

Goodman, L. 1974. Exploratory latent structure analysis using both identifiable and unidentifiable models. *Biometrika* 61:2:215–231.

Hatzivassiloglou, V., and McKeown, K. 1997. Predicting the semantic orientation of adjectives. In *ACL-EACL 1997*, 174–181.

Hatzivassiloglou, V., and Wiebe, J. 2000. Effects of adjective orientation and gradability on sentence subjectivity. In *18th International Conference on Computational Linguistics (COLING-2000)*.

Hovy, E. 1998. Combining and standardizing large-scale practical ontologies for machine translation and other uses. In *In Proc. 1st International conference on language resources and evaluation (LREC)*.

Kleinberg, J. 1998. Authoritative sources in a hyperlinked environment. In *Proc. ACL=SIAM Symposium on Discrete Algorithms*, 226–233.

Lee, L. 1999. Measures of distributional similarity. In *Proc. ACL '99*, 25–32.

Lin, D. 1994. Principar–an efficient, broad-coverage, principle-based parser. In *Proc. COLING '94*, 482–488.

Lin, D. 1998. Automatic retrieval and clustering of similar words. In *Proc. COLING-ACL '98*, 768–773.

Lyons, J. 1977. *Semantics, Volume 1*. Cambridge, MA: Cambridge University Press.

Mahesh, K., and Nirenburg, S. 1995. A situated ontology for practical nlp. In *Proc. Workshop on Basic Ontological Issues in Knowledge Sharing*. International Joint Conference on Artificial Intelligence (IJCAI-95), Aug. 19-20, 1995. Montreal, Canada.

Marcus, M.; Santorini; B.; and Marcinkiewicz, M. 1993. Building a large annotated corpus of English: The penn treebank. *Computational Linguistics* 19(2):313–330.

Miller, G. 1990. Wordnet: An on-line lexical database. *International Journal of Lexicography* 3(4):??–??

Procter, P. 1978. *Longman Dictionary of Contemporary English*. Addison Wesley Longman.

Rooth, M.; Riezler, S.; Prescher, D.; Carroll, G.; and Beil, F. 1999. Inducing a semantically annotated lexicon via em-based clustering. In *Proc. 37th Annual Meeting of the Association for Computational Linguistics (ACL-99)*, 104–111.

Sack, W. 1995. Representing and recognizing point of view. In *Proc. AAAI Fall Symposium on AI Applications in Knowledge Navigation and Retrieval*.

Spertus, E. 1997. Smokey: Automatic recognition of hostile messages. In *Proc. IAAI*.

Terveen, L.; Hill, W.; Amento, B.; McDonald, D.; and Creter, J. 1997. Building task-specific interfaces to high volume conversational data. In *Proc. CHI '97*, 226–233.

Wiebe, J.; Bruce, R.; and O'Hara, T. 1999. Development and use of a gold standard data set for subjectivity classifications. In *Proc. 37th Annual Meeting of the Assoc. for Computational Linguistics (ACL-99)*, 246–253. University of Maryland: ACL.

Wiebe, J. 1994. Tracking point of view in narrative. *Computational Linguistics* 20(2):233–287.