A Dictionary of Nonsubsective Adjectives

Neha Nayak, Mark Kowarsky, Gabor Angeli, Christopher D. Manning
Stanford University
Stanford, CA 94305
{nayakne, markak, angeli, manning}@stanford.edu

Abstract

Computational approaches to inference and information extraction often assume that adjective-noun compounds maintain all the relevant properties of the unmodified noun. A significant portion of nonsubsective adjectives violate this assumption. We present preliminary work towards a classifier for these adjectives. We also compile a comprehensive list of 60 nonsubsective adjectives including those used for training and those found by the classifiers.

1 Introduction

Many NLP tasks must reason about adjective-noun compounds. For instance, in inference tasks, many systems assume that a property of a noun holds for every associated adjective-noun compound; similarly, in information extraction, adjective-noun compounds are often taken as justification for the extraction of the noun. In such applications, it is convenient to assume that all such adjectives are subsective – that is, any instance denoted by the adjective-noun compound is an instance of the noun. However nonsubsective adjectives, such as *former*, *alleged*, or *counterfeit*, violate this assumption.

We present an expanded classification scheme for such *nonsubsective* adjectives aimed towards NLP applications. This includes both the traditional taxonomic classification, which is relevant for tasks like information extraction, as well as a classification based directly on maintaining validity for natural language inference tasks.

We then present 60 instances of nonsubsective adjectives. Some of these adjectives are collected from the literature, and others are the output of a high-recall classifier trained from statistics over a large corpus of text. A total of 15 of our adjectives are recovered from the classifier, with only minimal annotation effort. To the best of the authors' knowledge, this is the largest synthesis of nonsubsective adjectives in the literature. Finally, we present an analysis of the adjectives collected, including practical considerations for applications to inference and information extraction.

2 Related Work

We base our taxonomy on existing work in the literature (Chierchia and McConnell-Ginet, 2000; Kamp, 1975; Kamp and Partee, 1995), but extend it to include a more fine-grained division of the subclasses of adjectives which are problematic for NLP tasks.

(Amoia and Gardent, 2006) proposes a classification of English adjectives geared towards the task of RTE (Dagan et al., 2006). In addition to the denotation-based subclasses of (Kamp and Partee, 1995), they classify 300 English adjectives based on syntactic features and the kind of semantic opposition they participate in, as well as making note of entailments prompted by the morphology of the adjectives. Inference patterns were defined for each of the fine-grained adjective classes. (Amoia and Gardent, 2007) tested the inference rules developed in (Amoia and Gardent, 2006) on a test suite labeled for entailment. Our taxonomy focuses on denotation-based classification.

(Pustejovsky, 2013) examined the inference patterns licensed by plain non-subsective adjectives as defined by the four-class distinction of (Kamp and Partee, 1995), based on the lexical context in which the modification occurs. Structure-to-inference mappings were identified for four types of contexts of a nonsubsective adjective. In contrast, we focus on a larger set of adjectives independent of their context, and consider more general inference patterns.

(Boleda et al., 2012) explored the possibility of modeling modification by nonsubsective adjectives as a first-order phenomenon in vector space.

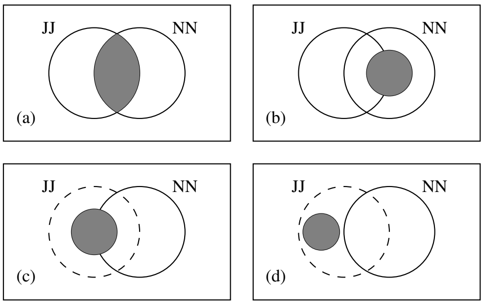

Figure 1: A visual representation of the classes of adjectives. The denotations of the noun NN and adjective JJ are given by hollow circles; strictly, non-subsective adjectives do not have denotations, their denotations here are given by broken circles; the denotation of the compound JJ NN is visually portrayed by the shaded circle. Figure (a) describes intersective adjectives; (b) describes strictly subsective adjectives, (c) describes plain nonsubsective adjectives, and (d) describes privative adjectives.

They showed that existing distributional models found it significantly more difficult to automatically model modification by a fixed set of nonsubsective adjectives. (Boleda et al., 2013) amended this setup to encompass a less restricted set of subsective adjectives, finding that subsective and nonsubsective adjectives have similar distributional behavior. This result suggests value in a detailed analysis of these adjectives not only for its intrinsic value, but also to help inform automatic methods for modeling and identifying them.

3 Theoretical Framework

We describe and motivate our categorization of adjectives, and introduce notation and terminology used throughout the paper.

3.1 Taxonomy

We let JJ stand for an adjective, and NN stand for an noun. The *denotation* of a phrase x, $[\![x]\!]$, is defined as the set of objects identified by the phrase. For example, $[\![cat]\!]$ is the set of cats, $[\![blue]\!]$ is the set of all blue things, and $[\![blue\ cat]\!]$ is the set of blue cats. The classical criterion for classifying adjectives characterizes the relationship between the denotations of the JJ NN phrase and the denotations of its constituents.

This is represented visually in Figure 1; this work focuses on the *plain nonsubsective* and *privative* adjectives in (c) and (d), respectively. Each of the classes is further described below.

Intersective The most common class of adjectives fall into the intersective category: Figure 1 (a). The denotation of the adjective-noun compound is the intersection of the denotations of its constituents. For example, a *blue box*, or *raggedy man*. Formally: $[\![\text{JJ NN}]\!] = [\![\text{JJ}]\!] \cap [\![\text{NN}]\!]$.

Subsective The second class of adjectives – Figure 1 (b) – are subsective adjectives. The denotation of the adjective-noun compound is a subset of the denotation of the noun, but is not necessarily a subset of the denotation of the adjective. For example, a *large thimble* (not necessarily large), or *cold star* (not necessarily cold). Formally: $[\![\text{JJ NN}]\!] \subseteq [\![\text{NN}]\!]$. Note that the intersective classification is a special case of subsective.[1]

Non-subsective The third class of adjectives – Figure 1 (c) and (d) – are the nonsubsective adjectives.[2] This class is the primary focus of this paper. The class is often subdivided into *plain nonsubsective* and *privative*.

The denotation of a noun modified by a nonsubsective adjective may still intersect with the denotation of the noun. For example, a *former* governor cannot be a governor, but an *alleged* criminal may be a criminal.

Adjectives for which the denotation of the adjective-noun compound is disjoint from the denotation of the noun are classified as *privative*.

Formally: $[\![\text{JJ NN}]\!] \cap [\![\text{NN}]\!] = \varnothing$ For example, *former*, *virtual*, and *fake*.

In contrast, the denotation of *plain nonsubsective* adjective compounds may intersect with the denotation of the noun: $[\![\text{JJ NN}]\!] \cap [\![\text{NN}]\!] \neq \varnothing$. For example, *alleged*, *possible*, and *unlikely*.

Additional Classes We introduce two subclasses of privative adjectives: those which are *counterfactual*, and those which exhibit a *temporal shift*. Counterfactual adjectives (*fake*, *mistaken*, etc.) constitute the more conventional class of privative adjectives; however, for many applications it is useful to distinguish whether an instance of the compound was ever or will ever be within the denotation of the noun. For example, *former*,

[1] *Extensional* is often used to describe subsective and intersective adjectives.

[2] *Intensional* is often used in the literature to describe nonsubsective adjectives.

and *future* appear in compounds describing objects that are not currently in the denotation of the noun. However, this does not hold for these objects at all points in time.

We note that the definitions above are a classification over senses of adjectives, rather than over types. For example, the sense of *apparent* which is synonymous with *visible* is intersective, while the sense synonymous with *ostensible* is plain non-subsective .

3.2 Categorization by Necessary Properties

We consider the set of properties that an object must have to belong to the denotation of some noun with certainty. We define these intrinsic properties of a noun NN in modal logic as the set of predicates P which necessarily hold over the noun:

$$\forall x.x \in [\![\text{NN}]\!] \rightarrow \Box P(x), \text{abbreviated as} \Box P(\text{NN})$$

For example, a gun has a necessary property of *shoots bullets*, and a refrigerator has a necessary property of *keeps things cold*.

We categorize adjectives based on the proportion of necessary properties that they preserve. Most subsective adjectives, including intersective adjectives, preserve all intrinsic properties of a noun :

$$\forall P \left[\Box P(\text{NN}) \rightarrow \Box P(\text{JJ NN})\right]$$

Certain nonsubsective adjectives, like *former*, preserve *most* intrinsic properties of a noun (*Most* in the table). For example, except for the property of being in office, a *former president* probably has many other properties in common with a *president*.

In contrast, a *fictional cat* is exempt from almost any particular attribute of a *cat*. (*None* in the table):

$$\neg\exists P \left[\Box P(\text{NN}) \rightarrow \Box P(\text{JJ NN})\right]$$

Furthermore, identifying subsective adjectives which do not preserve intrinsic properties is of interest. For instance, although an *erroneous attribution* is an attribution, it lacks the intrinsic property of attributing a work to its creator.

4 Data and Analysis

We compiled a list of 60 non-subsective adjectives from both prior work and a high-recall classifier. This list is presented in Table 1, along with relevant features of these adjectives for NLP tasks. We describe the sources for these adjectives, and expand on both the features in the table and the practical impact of these features on NLP applications in the later sections.

4.1 Data Sources

The list of adjectives proposed as nonsubsective was collected from three broad data sources: prior work, a high-recall classifier, and synonyms of known adjectives.

The adjectives from the literature were collected from (Partee, 2009), (Partee, 2010), (Boleda et al., 2012), (Boleda et al., 2013), and (Pustejovsky, 2013); in Table 1, these are denoted as P09, P10, B12, B13, and P13 respectively. Finally, we added synonyms of the known non-subsective adjectives to the list. In addition, we expanded the list by adding morphological variants of the known non-subsective adjectives; for example, *improbable* from *probable*.

4.2 List of Adjectives

We present our list of adjectives in Table 1. In addition to the adjective gloss, when available we include the WordNet (Miller, 1995) synsets of the senses which behave in a non-subsective way. The definition and source of the adjective are also provided.

The subclass of the adjective is then specified, according to the extended taxonomy in Section 3. *Modal* corresponds to adjectives that indicate uncertainty. *Temporal* indicates that $[\![\text{JJ NN}]\!]$ is not currently a subset of members of $[\![\text{NN}]\!]$, but is at some other time. The third class – *counterfactual* – affirms that an adjective-noun compound is in contradiction with being an instance of the noun.

Finally, the *Taxonomy* column denotes whether an adjective should be considered non-subsective using the taxonomic definition of the category. The *Properties* column, in turn, characterizes whether most or some of the fundamental properties of the noun necessarily hold for the adjective-noun compound. It's worth noting, as mentioned in Section 5, that some adjectives (e.g., spurious) appear subsective taxonomically but relax requirements for some or most fundamental properties of their associated noun.

Word	Syn	Definition	Source	Subclass	Properties	Word	Syn	Definition	Source	Subclass	Properties
alleged		declared but not proved	B13	Modal	Some	**anti-**		not in favor of (an action or proposal etc.)	Class	Cf.	None
believed	-	*accepted or regarded as true	P13	Modal	Some	**fabricated**		formed or conceived by the imagination	P10	Cf.	None
debatable	1,3	open to doubt or debate	Syn	Modal	Some	**fake**	2	fraudulent; having a misleading appearance	P10	Cf.	None
disputed		subject to disagreement and debate	P10	Modal	Some	**fictional**	2	formed or conceived by the imagination	Class	Cf.	None
dubious	1,2	fraught with uncertainty or doubt	Syn	Modal	Some	**fictitious**		formed or conceived by the imagination	P10	Cf.	None
hypothetical		based primarily on surmise rather than adequate evidence	B13	Modal	Some	**imaginary**		not based on fact; existing only in the imagination	P10	Cf.	None
impossible	1,2	not capable of occurring or being accomplished or dealt with	B13	Modal	Some	**mythical**		based on or told of in traditional stories; lacking factual basis or historical validity	P10	Cf.	None
improbable	1,2	not likely to be true or to occur or to have occurred	Pre	Modal	Some	**phony**		fraudulent; having a misleading appearance	Syn	Cf.	None
plausible		apparently reasonable and credible	Class	Modal	Some	**false**	1-7,9	not in accordance with the fact or reality or actuality	B12	Cf.	None
putative		purported; commonly put forth or accepted as true on inconclusive grounds	B13	Modal	Some	**artificial**	1,3	contrived by art rather than nature	B12	Cf.	Some
questionable		subject to question	P10	Modal	Some	**erroneous**†		containing or characterized by error	Class	Cf.	Some
so-called		doubtful or suspect	P13	Modal	Some	**mistaken**†		wrong in e.g. opinion or judgment	Class	Cf.	Some
supposed	2,3,4	mistakenly believed	P13	Modal	Some	**mock**		constituting a copy or imitation of something	B13	Cf.	Some
suspicious	2	not as expected	Class	Modal	Some	**pseudo-**		(often used in combination) not genuine but having the appearance of	Syn	Cf.	Some
theoretical	-	*relating to what is possible or imagined rather than to what is known to be true or real	B13	Modal	Some	**simulated**		not genuine or real; being an imitation of the genuine article	Class	Cf.	Some
uncertain	2,3,4	not established beyond doubt; still undecided or unknown	Class	Modal	Some	**spurious**†	1,2,4	"ostensibly valid	P10	Cf.	Some
unlikely	3	having a probability too low to inspire belief	Pre	Modal	Some	**unsuccessful**†		not successful; having failed or having an unfavorable outcome	Class	Cf.	Some
would-be		unfulfilled or frustrated in realizing an ambition	P10	Modal	Some	**counterfeit**		not genuine; imitating something superior	P10	Cf.	Most
doubtful	1	open to doubt or suspicion	P09	Modal	Some	**deputy**	4	a person appointed to represent or act on behalf of others	Class	Cf.	Most
apparent	2	appearing as such but not necessarily so	B12	Modal	Most	**faulty**†		having a defect	Class	Cf.	Most
arguable	1	capable of being supported by argument	P10	Modal	Most	**virtual**		being actually such in almost every respect	Syn	Cf.	Most
assumed	-	*thought to be true or probably true without knowing that it is true	Class	Modal	Most	**erstwhile**		belonging to some prior time	Class	Temp.	Most
likely		with considerable certainty; without much doubt	B13	Modal	Most	**ex-**	-	*one that formerly held a specified position or place	Syn	Temp.	Most
ostensible		appearing as such but not necessarily so	P10	Modal	Most	**expected**		considered likely or probable to happen or arrive	P09	Temp.	Most
possible	2	existing in possibility	B13	Modal	Most	**former**	2,3,4	belonging to some prior time	B13	Temp.	Most
potential		existing in possibility	B13	Modal	Most	**future**	1	yet to be or coming	B13	Temp.	Most
predicted		*said to happen or possible happen in the future	P10	Modal	Most	**historic**	1	belonging to the past; of what is important or famous in the past	Class	Temp.	Most
presumed		*thought to be true	B13	Modal	Most	**onetime**		belonging to some prior time	Syn	Temp.	Most
probable		likely but not certain to be or become true or real	B13	Modal	Most	**past**	2	of a person who has held and relinquished a position or office	B13	Temp.	Most
seeming		appearing as such but not necessarily so	Class	Modal	Most	**proposed**	-	set forth for acceptance or rejection	P09	Temp.	Most

Table 1: The list of non-subsective adjectives. The columns, from left to right, denote the adjective, the non-subsective word senses of the adjective, indexed by WordNet (3.1) synset, a definition of a non-subsective sense where available (adjectives where a non-subsective sense did not appear in WordNet are indicated by *), the source the adjective was collected from, the subclass of the adjective, and the proportion of fundamental properties that must hold for the modified noun. Adjectives that are taxonomically subsective are indicated by †.

4.3 Caveats

Although the set of adjectives indicated broadly describe a problematic subclass of adjectives, the presence of these adjectives alone is not sufficient to indicate non-subsective modification. Some adjectives are polysemous; for example *theoretical* (*theoretical physics* is *physics*) or *assumed* (an *assumed name* is a *name*), or occur in idiomatic collocations such a *false alarm* and *potential difference*.

The adjectives presented here are general in that they tend to be non-subsective regardless of the noun they modify. However, it is important to note that many intersective adjectives may have privative effects when co-occurring with certain nouns, as observed in (Partee, 2010). For example, *wooden* is usually intersective, but a *wooden lion*, while indeed being *wooden*, is not a *lion*.

5 Applications

5.1 Application for Information Extraction

An important challenge in information extraction is determining whether a sentence which appears to describe an extraction is reliable (Wiebe, 2000; Hatzivassiloglou and Wiebe, 2000). Modification by non-subsective adjectives, in turn, is often sufficient justification for questioning the reliability of the sentence.

For example, if we are given the sentence:

> *George Bush is the former president of the United States*,

the extraction claiming that George Bush is president may not be intended. Perhaps an even more dangerous example would be the case below, where a fictional entity should certainly not be extracted:

> *Fictional president Merkin Muffley, played by Peter Sellers, ...*

In these cases, the taxonomic classification is the most relevant feature from Table 1. To illustrate, despite a former president sharing many properties with a president, it is, for certain task descriptions, never a valid extraction. Conversely, despite a potential investor missing fundamental properties of an investor, we can nonetheless safely extract that Warren Buffet is an investor from him being a potential investor in a company.

5.2 Application for Inference

For inference tasks, the relevant feature of an adjective-noun compound is less its taxonomic classification directly so much as whether the truth of a predicate is maintained when a noun is modified by the adjective. For instance, the knowledge that presidents sign bills – corresponding to a predicate $\texttt{signs\_bills}(x)$ – should apply to *honorable* presidents but not to *former* presidents. Current systems for inference in natural language (MacCartney and Manning, 2009; Icard III, 2012) often consider all adjectives to be intersective.

For this application, the most relevant column of the table is the properties column. Entries denoted by *None* or *Some* are likely to lead to incorrect inferences. For example, a predicate applied to a *gun* would have a high probability of no longer holding for a *fake gun*. Likewise, a predicate applied to *intelligence* is much less likely to hold for *artificial intelligence*. The proportion of properties which must hold for the adjective-noun compound serves as a proxy for the degree to which it is risky to introduce the adjective during inference.

In contrast, many adjectives in the list are technically non-subsective, but could often safely be used in inference, because only a single or a few fundamental properties are not satisfied. These are denoted by *Most*. For instance, a *deputy department head*, *likely candidate*, or *former president*.

6 Experiments

6.1 Classification of adjectives

We defined a binary classification task, in which adjectives are classified as either *Subsective* (Corresponding to the classes *Intersective* and *Strictly subsective* defined in Section 3) or *Nonsubsective* (Corresponding to the classes *Privative* and *Plain non-subsective*). This classifier was motivated by a simple hypothesis - that subsective and nonsubsective adjectives differ in the nouns that they can modify. For example, nouns like *perpetrator*, which co-occur with intensional adjectives like *alleged* and *likely*, are likely to co-occur with other intensional adjectives.

We used three sets of examples, containing subsective and nonsubsective adjectives in the ratios 1:1, 1:10, 1:100. 30 of the known nonsubsective adjectives (Occurring in the table with any *Source* label besides *Class*) comprised the nonsubsective class. The subsective class was populated with subsective adjectives of comparable frequencies, using the (faulty) assumption that all adjectives not known to be nonsubsective were subsective.

Using adjective-noun bigrams from Wikipedia, we constructed a vector space model with adjectives as its elements, using their co-occurrence frequencies with nouns as features. A linear support vector machine (SVM) was used (Pedregosa and others, 2011; Fan et al., 2008), with the penalty for errors for each class set to be inversely proportional to their frequency in the training data. The co-occurrence matrix was weighted using PMI^2 (Bouma, 2009), and feature selection was performed using a model for Differential Expression (Robinson and others, 2010).

$$\mathrm{PMI}^2(A, N) = \log\left(\frac{\Pr(A \cap N)^2}{\Pr(A) \cdot \Pr(N)}\right)$$

This classifier performed poorly in all three cases. The accuracy did not exceed the majority baseline. Although this classifier did not perform well, observing the false positives in its results revealed many nonsubsective adjectives that were not used previously in the literature. In fact, one quarter of the set of adjectives we present originate from the classifier. These are denoted by *Class* in the *Source* column of Table 1.

6.2 Classification of adjective-noun pairs

The simplistic hypothesis described above does not adequately encapsulate the idea of nonsubsective modification. We implemented a more refined hypothesis - that subsectively modified nouns would be distributionally more similar to their unmodified counterparts than nonsubsectively modified nouns.

For example, we expect the difference between the distribution of contexts of *fake handbag*, and those of *handbag* to be greater than the difference between the corresponding distributions from *brown handbag* and *handbag*, as these involve nonsubsective and subsective modification, respectively. This model captures the differences such as that between the occurrences of *assumed* in *assumed culprit* and *assumed name*, as the choice of noun selects the sense of the adjective.

We assembled a set of frequent adjective-noun bigrams for each adjective in Table 1, as well as each subsective adjective. To quantify the difference between the distribution of these bigrams' contexts and that of the unmodified nouns, we used the following five measures of similarity.

$$\mathrm{KL}(\mathrm{P_{AN}}||\mathrm{P_N}) = \sum_x P_{AN}(x) log\left(\frac{P_{AN}(x)}{P_N(x)}\right)$$

$$\mathrm{KL}(\mathrm{P_N}||\mathrm{P_{AN}}) = \sum_x P_N(x) log\left(\frac{P_N(x)}{P_{AN}(x)}\right)$$

$$\mathrm{JS}(\mathrm{P_N}, \mathrm{P_{AN}}) = \frac{1}{2}\left(KL(P_{AN}||P_N) + KL(P_N||P_{AN})\right)$$

$$\mathrm{Cosine(N, AN)} = \frac{w_N \cdot w_{AN}}{||w_N||\,||w_{AN}||}$$

$$\mathrm{Jaccard(N, AN)} = \frac{||min(w_N, w_{AN})||_1}{||max(w_N, w_{AN})||_1}$$

x denotes a unique context. The features used were sentence-level bag-of-words contexts, n-gram contexts, and dependency paths. The experiments with bag-of-words contexts and n-gram contexts were repeated using vectors generated by word2vec (Mikolov et al., 2013a; Mikolov et al., 2013b). Decision tree classifiers (Pedregosa and others, 2011) were used. The highest F_1 score attained was 29%.

7 Conclusion

We have presented a synthesis of nonsubsective adjectives, and explored some relevant properties of these adjectives for NLP applications. We outlined some attempts at automatically detecting these adjectives and their properties using computational approaches, as well as identifying situations where a subsective adjective is nonsubsective in the context of a particular noun.

The task of identifying and characterising nonsubsective modification is important for inference and information extraction. Although the classifiers were unsuccessful, it is hoped that the list of adjectives accumulated in the course of this work will be useful for future work on this task.

References

Marilisa Amoia and Claire Gardent. 2006. Adjective based inference. In *Proceedings of the Workshop on Knowledge and Reasoning for Language Processing*. Association for Computational Linguistics.

Marilisa Amoia and Claire Gardent. 2007. A first order semantic approach to adjectival inference. In *Proceedings of the ACL-PASCAL Workshop on Textual Entailment and Paraphrasing*. Association for Computational Linguistics.

Gemma Boleda, Eva Maria Vecchi, Miquel Cornudella, and Louise McNally. 2012. First-order vs. higher-order modification in distributional semantics. In

EMNLP. Association for Computational Linguistics.

Gemma Boleda, Marco Baroni, Louise McNally, and Nghia Pham. 2013. Intensionality was only alleged: On adjective-noun composition in distributional semantics. In *Proceedings of IWCS*.

Gerlof Bouma. 2009. Normalized (pointwise) mutual information in collocation extraction. *Proceedings of GSCL*, pages 31–40.

Gennaro Chierchia and Sally McConnell-Ginet. 2000. *Meaning and grammar: An introduction to semantics*. MIT press.

Ido Dagan, Oren Glickman, and Bernardo Magnini. 2006. The pascal recognising textual entailment challenge. In *Machine Learning Challenges. Evaluating Predictive Uncertainty, Visual Object Classification, and Recognising Tectual Entailment*, pages 177–190. Springer.

Rong-En Fan, Kai-Wei Chang, Cho-Jui Hsieh, Xiang-Rui Wang, and Chih-Jen Lin. 2008. Liblinear: A library for large linear classification. *The Journal of Machine Learning Research*, 9:1871–1874.

Vasileios Hatzivassiloglou and Janyce M Wiebe. 2000. Effects of adjective orientation and gradability on sentence subjectivity. In *Proceedings of the 18th conference on Computational linguistics-Volume 1*, pages 299–305. Association for Computational Linguistics.

Thomas F Icard III. 2012. Inclusion and exclusion in natural language. *Studia Logica*.

Hans Kamp and Barbara Partee. 1995. Prototype theory and compositionality. *Cognition*, 57(2):129–191.

Hans Kamp. 1975. Two theories about adjectives. *Formal semantics of natural language*, pages 123–155.

Bill MacCartney and Christopher D Manning. 2009. An extended model of natural logic. In *Proceedings of the eighth international conference on computational semantics*. Association for Computational Linguistics.

Tomas Mikolov, Kai Chen, Greg Corrado, and Jeffrey Dean. 2013a. Efficient estimation of word representations in vector space. *arXiv preprint arXiv:1301.3781*.

Tomas Mikolov, Ilya Sutskever, Kai Chen, Greg S Corrado, and Jeff Dean. 2013b. Distributed representations of words and phrases and their compositionality. In *Advances in Neural Information Processing Systems*, pages 3111–3119.

George A. Miller. 1995. WordNet: a lexical database for English. *Communications of the ACM*.

Barbara H Partee. 2009. Formal semantics, lexical semantics and compositionality: The puzzle of privative adjectives. *Philologia*, (7):11–24.

Barbara H Partee. 2010. Privative adjectives: subsective plus coercion. *Presuppositions and Discourse: Essays Offered to Hans Kamp*.

F. Pedregosa et al. 2011. Scikit-learn: Machine learning in Python. *Journal of Machine Learning Research*, 12:2825–2830.

James Pustejovsky. 2013. Inference patterns with intensional adjectives. In *Workshop on Interoperable Semantic Annotation*.

Mark D Robinson et al. 2010. edger: a bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics*, 26(1):139–140.

Janyce Wiebe. 2000. Learning subjective adjectives from corpora. In *AAAI/IAAI*, pages 735–740.