A Test Suite for Inference Involving Adjectives

Marilisa Amoia*, **Claire Gardent**†

* University of Saarland
Saarbrücken, Germany
amoia@coli.uni-saarland.de

†CNRS/Loria
Campus Scientifique BP 239, Vandoeuvre-les-Nancy, France
gardent@loria.fr

Abstract

Recently, most of the research in NLP has concentrated on the creation of applications coping with textual entailment. However, there still exist very few resources for the evaluation of such applications. We argue that the reason for this resides not only in the novelty of the research field but also and mainly in the difficulty of defining the linguistic phenomena which are responsible for inference. As the TSNLP project has shown test suites provide optimal diagnostic and evaluation tools for NLP applications, as contrary to text corpora they provide a deep insight in the linguistic phenomena allowing control over the data. Thus in this paper, we present a test suite specifically developed for studying inference problems shown by English adjectives. The construction of the test suite is based on the deep linguistic analysis and following classification of entailment patterns of adjectives and follows the TSNLP guidelines on linguistic databases providing a clear coverage, systematic annotation of inference tasks, large reusability and simple maintenance. With the design of this test suite we aim at creating a resource supporting the evaluation of computational systems handling natural language inference and in particular at providing a benchmark against which to evaluate and compare existing semantic analysers.

1. Introduction

In the last years, due to the fact that the web has become the major source of information, research in question answering, information retrieval and information extraction has focused on developing systems which can handle natural language inference. Such applications should indeed be able to recognize that some text H contains information which is implicitly given in another text T (entailment recognition task) and/or that two texts convey the same information (paraphrase recognition task).

(1) a. *The bones fossilised*

b. *The bones got preserved in stone*

(2) a. *Janet broke the window*

b. *The window is broken*

Given the great economic importance of information extraction and recognition techniques, most systems focus on providing wide coverage and robustness thus relying on statistical methods.

As the Pascal Recognizing Textual Entailment RTE Challenge (Dagan et al., 2006) has shown, semantic construction and reasoning are two crucial components in enhancing the quality and accuracy of such NLP systems. However, there exist still very scarce resources which provide test data focusing on entailment problems. The collections of test data appeared till now (see RTE data sets, the Microsoft Paraphrase Research Corpus (Dolan et al., 2004), Text Retrieval Conference TREC data), created to evaluate systems for wide coverage and robustness, do not satisfy, in our opinion, other important requirements necessary to test systems handling inference. First, none of these collections is annotated for inference tasks, so that it is not clear what their linguistic coverage is. Second, the examples are often taken from newspaper articles so that they present a quite big syntactic complexity and are often difficult to be used for the evaluation of symbolic based semantic analysers. Third, they use a quite loose notion of semantic equivalence and entailment so that the examples are often difficult to judge.

On the contrary, we think that resources which give a deeper insight in the linguistic phenomena which are responsible for inference may help in enhancing the ability of applications to cope with it. Thus, in this paper, we address the entailment problem from a linguistic based perspective and present a test suite which focuses on a specific linguistic task, namely adjectival inference and address this issue deeply. Moreover, we use a well defined notion of entailment as we aim at providing a resource for also evaluating deep semantic analysers based on symbolic methods.

Thus, the test suite presented in this paper includes a systematic classification of adjectival inferential tasks and semantic annotations for adjectives based on WordNet and on the semantic classification of English adjectives proposed by (Amoia and Gardent, 2006).

This paper is structured as follows, we first define the notion of entailment we presuppose, then we describe the linguistic task we focus on and finally give some details of the realisation of the test suite.

2. Developing a Test Suite for Adjectival Inference Problems

In collecting the test data, we have followed the TSNLP (Balkan et al., 1994) guidelines for the development of linguistic test suites, so that this work meets the requirements of systematicity, neutral vocabulary and well-founded approach to test positive and negative cases. In the following we first define the notion of entailment we presuppose, then we describe linguistic task, lexical coverage and implementation of the test suite.

2.1. The Entailment Recognition Task

The idea behind the construction of this test suite is to illustrate the semantic and syntactic behavior of adjectives and their morphologically related verbs and nouns with respect to textual entailment. Thus, the test suite is a collection of sentence pairs (S1/S2) each illustrating a particular entailment problem: the first sentence in the pair (S1) can be recognized as entailing the second one (S2) if and only if the right type of inference (i.e. syntactic, semantic, lexical semantic or morphoderivational) is performed. The notion of textual entailment we use corresponds to the notion of logic entailment between the representations of the two texts:

Given two texts $\mathcal{T}_1$ and $\mathcal{T}_2$, it holds that $\mathcal{T}_1$ *entails* $\mathcal{T}_2$ iff: $\Phi(\mathcal{T}_1) \models \Phi(\mathcal{T}_2)$,

where $\Phi(\mathcal{T}_i)$ corresponds to the logic representation of the text T_i.

2.2. Linguistic Task

The construction of the test suite focuses on collecting specific classes of inference problems for English adjectives. In order to define the set of such inference problems, we build on (Amoia and Gardent, 2006; Amoia and Gardent, 2007) who have shown that in order to correctly predict adjectival inferential patterns it is important to consider the fine interplay between the different properties of adjectives which range from syntax and semantics to lexical semantics and morphoderivational properties. Thus, we have first individuated a set of general properties of adjectives by relaying on linguistic works on adjectives. Namely, we have merged together the syntactic properties of adjectives proposed by (Quirk et al., 1985), (Huddleston, 1984) and (Vendler, 1963) with the semantic properties proposed in (Chierchia and Connell-Ginet, 1990) and the model theoretic properties proposed by (Kamp and Partee, 1995) and (Keenan, 1987).
Then, we have extracted the inferential patterns which originate from these properties, thus obtaining a set of about 40 inferential patterns of adjectives originating from different sources such as syntax, model theoretic semantics, lexical semantics and derivational morphology. In the following, we describe in detail the patterns considered in the test suite.

2.2.1. Syntactic Patterns

The set of inference patterns with syntactic source we consider in the test suite includes the following syntactic alternations describing paraphrastic patterns:

P1: Predicative/Attributive Construction
`N is Adj ↔ This is Adj N`

(3) *This is a red table ↔ This table is red*

P2: For-Construction
`This is Adj N ↔ This is Adj for an N`

(4) *Jerry is a big mouse ↔ Jerry is big for a mouse*

P3: As-Construction
`This is Adj N ↔ This is Adj as an N`

(5) *John is a good cook ↔ John is good as a cook*

Furthermore, we consider adjectival constructions with clausal complement (SC) such as object embedding, subject embedding, easy/tough constructions and the inferential patterns they originate. All adjectives allowing subject embedding, for example, participate in the It-extraposition paraphrastic pattern.

P4: It-Extraposition
`It is Adj SC ↔ SC is Adj`

(6) *It is possible that it will rain tomorrow ↔ That it will rain tomorrow is possible*

As shown in (Arnold, 1989), some adjectives which allow subject embedding participate in the Of-PP paraphrastic pattern some others in the For-PP paraphrastic pattern , i.e. constructions in which the modified noun appears as a PP argument of the adjective.

P5: Of-Construction
`N is Adj SC ↔ It is Adj of N SC`

(7) *John is stupid to take this job ↔ It is stupid of John to take this job*

P6: For-Construction
`N is Adj SC ↔ SC is Adj for N`

(8) *I'm sad to leave ↔ To leave is sad for me*

As shown in (Flickinger and Nerbonne, 1992), some adjectives which allow subject embedding can participate in paraphrastic patterns called Easy/Though constructions, i.e. constructions in which the modified noun appears as a non-subject complement of the SC verb.

P7: Easy-Construction I
`N is Adj SC ↔ It is Adj SC`

(9) *John is easy to talk to ↔ It is easy to talk to John*

P8: Easy-Construction II
`N is Adj for-PP SC ↔ It is Adj of N SC`

(10) *John is easy for Mary to talk to ↔ It is easy for Mary to talk to John*

2.2.2. Lexical Semantics

By considering lexical semantic properties, other entailment patterns can be generated. For example, the different behaviour shown by adjectives with respect to their antonyms (Cruse, 1986) originates different entailment relations.

P9: Binary Antonymic Relations
`N is not A ↔ N is AntonymOf(A)`

(11) *The dishcloth is not wet ↔ The dishcloth is dry*

P10: Contrary Antonymic Relations I
`N is A` $\rightarrow$ `N is not AntonymOf(A)`

(12) *The mouse is small → The mouse is not big*

P11: Contrary Antonymic Relations II
`N is not A` $\not\rightarrow$ `N is AntonymOf(A)`

(13) *The mouse is not small ↛ The mouse is big*

Hyponymy between adjectival items is also a productive source of inference.

P12: Adjective Hyponymy
`N is Hypo(A)` $\rightarrow$ `N is A`

(14) *He is minuscule → He is small*

Other interesting inferential patterns can be captured by analysing whether the property expressed by the adjective is inherited by the hyponyms of the modified noun or not. Such patterns are applied by Kennedy (1998) to gradable adjectives to individuate logical polarity.

P13: Noun Hyponymy I
`This is Adj HypoOf(N)` $\rightarrow$ `This is Adj N`

(15) *X is a red table → X is a red object*

(16) *John is a civil lawyer ↛ John is a civil man*

(17) *This is a counterfeit diamond ↛ This is a counterfeit object*

(18) *He is a fictitious hero → He is a fictitious person*

(19) *John is the alleged strangler → John is the alleged murderer*

P14: Noun Hyponymy II
`This is Adj N` $\rightarrow$ `This is Adj HypoOf(N)`

(20) *John is a dangerous man → John is a dangerous husband*

P15: SC Hyponymy I
`It is Adj SC` $\rightarrow$ `It is Adj Hypo(SC)`

(21) *It is dangerous to drive in Rome → It is dangerous to drive fast in Rome*

(22) *It is safe to drive in Rome ↛ It is safe to drive fast in Rome*

P16: SC Hyponymy II
`It is Adj Hypo(SC)` $\rightarrow$ `It is Adj SC`

(23) *It is dangerous to drive fast in Rome ↛ It is dangerous to drive in Rome*

(24) *It is safe to drive fast in Rome → It is safe to drive in Rome*

By considering the taxonomical category of the adjective (that often can be extracted from WordNet) we obtain the following pattern.

P17: Taxonomical Category
`N is Adj` $\leftrightarrow$ `N has a Adj TaxoCat(Adj)`

(25) *This table is red ↔ This table has a red color*

(26) *This mouse is big ↔ This mouse has a big size*

(27) *This man is happy ↔ This man is in a happy mental state*

(28) *This book is good ↔ This book has a good quality*

2.2.3. Derivational Morphology

Building on (Vendler, 1963; Vendler, 1968) and (Quirk et al., 1985), we have collected entailment patterns which have derivational morphology as source. We use the following notational convention:

N the noun modified by the adjective,
Av an adjective A which is morphologically related to the verb V,
An an adjective A which is morphologically related to the noun N,
Aadv an adjective A which is morphologically related to the adverb ADV.
Va represents a verb V which is morphologically related to the adjective A,
ADVa an adverb ADV which is morphologically related to the adjective A.

The adjective-verb alternations describe constructions in which the modified noun N becomes the subject or the object of the morphologically related verb.

P18: Adjective-Verb Alternation I
`N is Av` $\leftrightarrow$ `N V`

(29) *John is asleep ↔ John sleeps*

P19: Adjective-Verb Alternation II
`N is Av` $\leftrightarrow$ `It is possible to V N`

(30) *This fungus is edible ↔ It is possible to eat this fungus*

P20: Adjective-Verb Alternation III
`N is Av Prep N1` $\leftrightarrow$ `N V N1`

(31) *This film is interesting for me ↔ This film interests me*

P21: Adjective-Verb Alternation IV
`N1 Va N` $\rightarrow$ `N is A`

(32) *John has opened the door → The door is open*

P22: Adjective-Verb Alternation V
`N1 is An2 Nv` $\leftrightarrow$ `N1 V N2`

(33) *He is the provincial governor ↔ He governs the province*

P23: Adjective-Verb Alternation VI
`N is ADV Av ↔ N V ADV`

(34) *John is deeply asleep* ↔ *John sleeps deeply*

Adjective-noun alternations describe constructions in which the adjective is substituted with a morphologically related noun.

P24: Adjective-ThetaRole_Noun Alternation
`N is An1 ↔ N is N1`

(35) *John is absent* ↔ *John is the absentee*

P25: Adjective-Event_Noun Alternation
`N is ADVa2 An1 ↔ N's N1 is A2`

(36) *John is deeply asleep* ↔ *John's sleep is deep*

P26: Adjective-NonEvent_Noun Alternation
`N is An1 ↔ N's N1`

(37) *John is polite* ↔ *John's politeness*

P27: Adjective-CategorialNoun Alternation
`This is a An1 N ↔ This N is a N1`

(38) *This is an Italian lawyer* ↔ *This lawyer is an Italian*

The relational adjective-noun alternations represent a set of inferential patterns which, as described in (Levi, 1978), differ for the particular relation Rel denoted by the adjective and syntactically realised as a different preposition in the paraphrase.

P28: Relational Adjective-Noun Alternation I
`This is An1 N ↔ This N is Rel(about) N`

(39) *This is a gastronomical dictionary* ↔ *This is a dictionary about gastronomy*

P29: Relational Adjective-Noun Alternation II
`This is An1 N ↔ This N is Rel(from) N`

(40) *They are rural visitors* ↔ *They are visitors from the country*

P30: Relational Adjective-Noun Alternation III
`This is An1 N ↔ This N is Rel(made of) N`

(41) *This is a wooden table* ↔ *This table is made of wood*

P31: Adjective-Adverb Alternation
`N1's Nv is Aadv ↔ N1 V ADV`

(42) *John's smile was cruel* ↔ *John smiled cruelly*

Constructions in which the modified noun is substituted with a prepositional phrase containing the verb implied by the noun. Different prepositions will generate different inferential patterns.

P32: (43) `N is A Nv↔ N is A at V-ing`
He is a good cook ↔ *He is good at cooking*

P33: (44) `N is A Nv↔ N is good for V-ing`
It is a good meal ↔ *It is good for eating*

1- `Dimension`	*big, large, little, small, ...*
2- `Physical property`	
2.1- `Sense`	*bitter, sweet, ...*
2.2- `Consistency`	*hard, soft, ...*
2.3- `Texture`	*rough, smooth, scaly, ...*
2.4- `Temperature`	*warm, cool, tepid, ...*
2.5- `Edibility`	*ripe, raw, cooked, ...*
2.6- `Substantiality`	*hollow, full, thick, ...*
2.7- `Configuration`	*sharp, broken, whole, ...*
3- `Speed`	*fast, quick, slow, ...*
4- `Age`	*new, young, old, ...*
5- `Color`	*red, blue, black, ...*
6- `Value`	*good, bad, perfect, ...*
7- `Difficulty`	*easy, difficult, ...*
8- `Qualification`	
8.1- `Definite`	*probable, ...*
8.2- `Possible`	*possible, ...*
8.3- `Usual`	*usual*
8.4- `Likely`	*likely*
8.5- `Sure`	*sure*
8.6- `Correct`	*appropriate*
9- `Human Propensity`	
9.1- `Mental State`	
9.1.1- `Fond`	*fond*
9.1.2- `Angry`	*jealous, angry, ...*
9.1.3- `Happy`	*anxious, happy, ...*
9.1.4- `Unsure`	*certain*
9.1.5- `Eager`	*eager, ready*
9.1.6- `Clever`	*clever, stupid, generous*
9.2- `Physical State`	*weak, sore, thirsty, ...*
9.3- `Behaviour`	*wild, funny, ...*
10- `Similarity`	*similar, different, ...*

Figure 1: Taxonomy of Adjectives

2.2.4. Model Theoretic Semantics

For model theoretic properties of adjectives, we relay on (Kamp and Partee, 1995) and (Keenan, 1987) which describe the semantics of adjective-noun combinations, i.e. the semantics of the attributive use of adjectives. Thus, we take the notion of *inherence* from (Quirk et al., 1985) and (Keenan, 1987), to describe inferential patterns in which the individual denotating the modified noun has the property expressed by the adjective. We use the notion of *subsectivity* and *privativity* described in (Kamp and Partee, 1995) respectively to describe patterns in which the adjective-noun combination allow to infer the noun property or its negation. The described properties originate the following inferential patterns.

P34: Inherence
`This is Adj N → This is Adj`

(45) *X is a red table* → *X is red*

(46) *John is a mechanical engineer* $\not\rightarrow$ *John is mechanical*

(47) *John is stupid to take this job* $\not\rightarrow$ *John is stupid*

P35: Subsectivity
`This is Adj N` $\rightarrow$ `This is N`

(48) *X is a red table* $\rightarrow$ *X is a table*

(49) *This is a counterfeit diamond* $\not\rightarrow$ *This is a diamond*

(50) *John is an alleged murderer* $\not\rightarrow$ *John is a murderer*

P36: Privativity
`This is Adj N` $\rightarrow$ `This is ¬N`

(51) *This is a counterfeit diamond* $\rightarrow$ *This is not a diamond*

(52) *This is an oval table* $\not\rightarrow$ *This is not a table*

(53) *John is an alleged murderer* $\not\rightarrow$ *John is not a murderer*

2.3. Lexical Coverage

One feature of adjectives which makes their analysis and classification difficult is their polysemy. Adjectives, in fact, can have different interpretations depending on the particular context in which they are uttered. As an example, consider the sentences below which show the polysemy of the adjective *heavy*.

(54) a. *This bag is heavy*

b. *John is a heavy smoker*

It is clear that *heavy* in (54a) has a dimensional meaning, while *heavy* in (54b) is a quality adjective. In order to cope with this problem, we define an adjectival item as corresponding to a WordNet (Fellbaum, 1998) sense of an adjective, i.e. to the WordNet reading corresponding to the interpretation of the adjective in the given example sentence. This choice is also motivated by the possibility to access the linguistic knowledge encoded in WordNet (antonyms, hyponyms, hyperonyms, etc.) and by the wide usage of this resource in NLP applications.
In order to collect a domain independent set of adjectival items we have proceeded as follows. Starting by the taxonomy of adjectives based on (Dixon, 1991) shown in Figure 1, we have choosen for each ontological category a set of items so as to obtain a sample containing adjectives displaying all different syntactic (e.g. adjectives which can be used predicatively and attributively, adjectives which allow only postnominal or attributive or predicative use), semantic (e.g. intersective, subsective, privative, plain nonsubsective adjectives), morphoderivational (e.g. denominal, deverbal, numeral adjectives) properties considered in the linguistic task. This initial sample was further expanded with synonyms, similar words, hyponyms and antonyms taken from WordNet. Thus, the final sample includes about 500 adjectival items.


```
<PAIR id=”1” entailment=”TRUE” inferencePattern=”P21”>
    <S1> <EXAMPLE>The dog frightened the child.</EXAMPLE>
        <SYNTAX>
            <N> dog </N>
            <V> frighten </V>
            <N> child </N>
        </SYNTAX> </S1>
    <S2> <EXAMPLE>The child is afraid.</EXAMPLE>
        <SYNTAX>
            <N> child </N>
            <COP/ >
            <ADJ wsn=”1” adjClass=”Ip2” > afraid </ADJ>
        </SYNTAX> </S2> </PAIR>
<PAIR id=”2” entailment=”FALSE” inferencePattern=”P11”>
    <S1> <EXAMPLE>This is not a rectangular table.</EXAMPLE>
        <SYNTAX>
            <N> this </N>
            <COP/ >
            <NEG/ >
            <ADJ wsn=”1” adjClass=”Ipa1” > rectangular </ADJ>
            <N> table </N>
        </SYNTAX> </S1>
    <S2> <EXAMPLE>This is a round table.</EXAMPLE>
        <SYNTAX>
            <N> this </N>
            <COP/ >
            <ADJ wsn=”1” adjClass=”Ipa1” > round </ADJ>
            <N> table </N>
        </SYNTAX> </S2></PAIR>
<PAIR id=”3” entailment=”TRUE” inferencePattern=”P13”>
    <S1> <EXAMPLE>John is a fictitious friend.</EXAMPLE>
        <SYNTAX>
            <N> John </N>
            <COP/ >
            <ADJ wsn=”1” adjClass=”PRpa1” > fictitious </ADJ>
            <N> friend </N>
        </SYNTAX> </S1>
    <S2> <EXAMPLE>John is a fictitious person.</EXAMPLE>
        <SYNTAX>
            <N> John </N>
            <COP/ >
            <ADJ wsn=”1” adjClass=”PRpa1” > fictitious </ADJ>
            <N> person </N>
        </SYNTAX> </S2></PAIR>
<PAIR id=”3” entailment=”FALSE” inferencePattern=”P13”>
    <S1> <EXAMPLE>John is a false doctor.</EXAMPLE>
        <SYNTAX>
            <N> John </N>
            <COP/ >
            <ADJ wsn=”6” adjClass=”PRpa2” > false </ADJ>
            <N> doctor </N>
        </SYNTAX> </S1>
    <S2> <EXAMPLE>John is a false person.</EXAMPLE>
        <SYNTAX>
            <N> John </N>
            <COP>
            <ADJ wsn=”6” adjClass=”PRpa2” > false </ADJ>
            <N> person </N>
        </SYNTAX> </S2></PAIR>
<PAIR id=”4” entailment=”FALSE” inferencePattern=”P35”>
    <S1> <EXAMPLE>John is an alleged murderer.</EXAMPLE>
        <SYNTAX>
            <N> John </N>
            <COP/ >
            <ADJ wsn=”2” adjClass=”PlNS3” > alleged </ADJ>
            <N> murderer </N>
        </SYNTAX> </S1>
    <S2> <EXAMPLE>John is a murderer.</EXAMPLE>
        <SYNTAX>
            <N> John </N>
            <COP>
            <N> murderer </N>
        </SYNTAX> </S2></PAIR>
<PAIR id=”5” entailment=”TRUE” inferencePattern=”P13”>
    <S1> <EXAMPLE>John is the alleged strangler.</EXAMPLE>
        <SYNTAX>
            <N> John </N>
            <COP/ >
            <ADJ wsn=”2” adjClass=”PlNS3” > alleged </ADJ>
            <N> strangler </N>
        </SYNTAX> </S1>
    <S2> <EXAMPLE>John is the alleged murderer.</EXAMPLE>
        <SYNTAX>
            <N> John </N>
            <COP>
            <ADJ wsn=”2” adjClass=”PlNS3” > alleged </ADJ>
            <N> murderer </N>
        </SYNTAX> </S2></PAIR>
```


Figure 2: An example of corpus annotation

2.4. Implementation

The test suite[1] contains a set of about 3000 sentence pairs which illustrate particular inference problems of adjectives, i.e. show inference patterns in which semantic, syntactic and morphoderivational criteria are the source of inference. In order to limit the problem, the sentence pairs contain texts with little syntactic complexity. So for example, many sentences follow the pattern NP V NP, where the verb is often the copula. These sentences were taken in some part from the literature on adjectives, in some part are hand coded, but mostly come from texts found on the Web and simplified at need. The example sentences have been created by generating for each adjectival item sentences representing all inferential patterns possible for that adjective. We have tried to consider an equal number of positive and negative cases of entailment.

Figure 2 shows an example of annotation. The test suite is encoded as an XML file. Each item in the test suite describes a sentence pair S1/S2 and includes

- a judgment about the truth of the entailment between the sentences in the pair. Thus, the attribute `entailment` has the values TRUE and FALSE, to respectively tag true and false entailment between the sentences S1 and S2 and TRUEDouble and FALSE-Double to signalise true and false cases of paraphrases

- a description of the type of inference problem shown in the sentence pair. The attribute `inferencePattern` has as a value the name of one of the patterns described in this paper. So for example, subsective patterns are annotated with `inferencePattern=P35`.

Moreover, each adjective is annotated with the WordNet sense (`wns`) and with the semantic class (`adjClass`) to which it corresponds. For the semantic class assignment, we use a refined version of the semantic classification of adjectives presented in (Amoia and Gardent, 2006), which includes about 30 semantic adjectival classes.

We would like to stress that the information with which the adjectival items are tagged, i.e. WordNet sense and adjectival class, are semantic information which can help reconstruct the meaning of the sentences thus enabling the automatic judgment of whether the entailment between the sentences in a given pair holds or not. The adjectival class assignment infact, points to a semantic representation of the adjective which is first order and compositional as described in (Amoia and Gardent, 2007).

3. Conclusion

In this paper we presented a test suite specifically created to study the inferential behavior of English adjectives. We hope it may serve as a resource for the evaluation of systems handling with natural language inference. With the construction of this test suite we want to open the way for the creation of resources which give a deeper insight in the linguistic phenomena which are responsible for inference.

We are aware of the limits of the test items included in our test suite, as we have considered only base cases of entailment. In the future we want to concentrate on the extension of the test sample by increasing the complexity of the test items to include cases which results from the combination of simpler ones.

4. References

Marilisa Amoia and Claire Gardent. 2006. Adjective based inference. In *Proceedings of KRAQ'06 (Knowledge and Reasoning for Answering Questions)*, Trento, Italy.

Marilisa Amoia and Claire Gardent. 2007. A first order semantic approach to adjectival inference. In *Proceedings of ACL-Pascal Workshop on Textual Entailment and Paraphrasing*, Prague, Czech Republic.

D. Arnold. 1989. *Theoretical and descriptive issues in machine Translation, PhD dissertation*. University of Essex.

Lorna Balkan, Siety Meijer, Doug Arnold, Eva Dauphin, Dominique Estival, Kirsten Falkedal, Sabine Lehmann, and Sylvie Regnier-Prost. 1994. Test suite design: Guidelines and methodology. In *Report to LRE 62-089*. University of Essex.

G. Chierchia and S. Mc Connell-Ginet. 1990. *Meaning and Grammar: An Introduction to Semantics*. The MIT Press, Cambridge, MA.

DA. Cruse. 1986. *Lexical Semantics*. Cambridge University Press.

Ido Dagan, Oren Glickman, and Bernardo Magnini. 2006. *The PASCAL Recognising Textual Entailment Challenge*, volume 3944.

R.M.W. Dixon. 1991. *A new approach to English grammar on semantic principles*. Clarendon Press, Oxford.

W. B. Dolan, C. Quirk, and C. Brockett. 2004. *Unsupervised Construction of Large Paraphrase Corpora: Exploiting Massively Parallel News Sources*. Geneva, Switzerland.

C. Fellbaum. 1998. Wordnet an electronic lexical database. Cambridge, MA, The MIT Press.

Dan Flickinger and John Nerbonne. 1992. Inheritance and complementation: A case study of *Easy* adjectives and related nouns. *Computational Linguistics*, 18(3):269–310.

Rodney Huddleston. 1984. *Introduction to the Grammar of English, Cambridge Textbooks in English*. Cambridge University Press.

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

Edward Keenan. 1987. Facing the truth: Some advantages of direct interpretation. *Universal Grammar*, (15 Essays):459–499.

Christofer Kennedy. 1998. *On the Monotonicity of Polar Adjectives*.

Judith Levi. 1978. *The syntax and semantics of complex nominals*. Academic Press, New York.

R. Quirk, S. Greenbaum, G. Leech, and J. Svartvik. 1985. *A Comprehensive Grammar of the English Language*. Longman.

[1]The test suite presented in this paper is available at http://www.coli.uni-sb.de/∼amoia/project/adj-TS

Zeno Vendler. 1963. The grammar of goodness. *The Philosophical Review*, 72(4):446–465.

Zeno Vendler. 1968. *Adjectives and Nominalization*. The Hague: Mouton.