Linguistic-Based Computational Treatment of Textual Entailment Recognition

Dissertation
zur Erlangung des akademischen Grades eines
Doktors der Philosophie
der Philosophischen Fakultäten
der Universität des Saarlandes

vorgelegt von Marilisa Amoia
aus Bari, Italien

Saarbrücken, 2009

Der Dekan:	Prof. Dr. R. Brünker
Berichterstatter:	Prof. Dr. M. Pinkal
	Dr. C. Gardent
Tag der letzten Prüfungsleistung:	7.11.2008

Abstract

In this thesis, I investigate how lexical resources based on the organisation of lexical knowledge in classes which share common (syntactic, semantic, etc.) features support natural language processing and in particular symbolic recognition of textual entailment. First, I present a robust and wide coverage approach to lexico-structural verb paraphrase recognition based on Levin's (1993) classification of English verbs. Then, I show that by extending Levin's framework to general inference patterns, a classification of English adjectives can be obtained that compared with previous approaches, provides a more fine grained semantic characterisation of their inferential properties. Further, I develop a compositional semantic framework to assign a semantic representation to adjectives based on an ontologically promiscuous approach (Hobbs, 1985) and thereby supporting first order inference for all types of adjectives including extensional ones. Finally, I present a test suite for adjectival inference I developed as a resource for the evaluation of computational systems handling natural language inference.

Résumé

Dans cette thèse, j'étudie la manière dont les ressources lexicales basées sur l'organisation de la connaissance lexicale dans des classes qui partagent des propriétés communes (syntactiques, sémantiques, etc.) permettent le traitement automatique de la langue naturelle et en particulier la reconnaissance symbolique d'implications textuelles.
Tout d'abord, je présente une approche robuste et à large couverture sur la reconnaissance de paraphrases verbales lexico-structurelle basée sur la classification de verbes anglais par Levin (1993). Puis, je montre qu'en étendant le cadre proposé par Levin pour traiter les modèles d'inférence généraux, on obtient une classification d'adjectifs anglais qui, comparée à des approches antérieures, propose une caractérisation sémantique à grain plus fin de leurs propriétés déductives. De plus, je développe un cadre sémantique compositionnel pour assigner à des adjectifs une représentation sémantique sur la base d'une approche ontologiquement variée (Hobbs, 1985) et qui permet ainsi l'inférence de premier ordre pour tous les types d'adjectifs, y compris les adjectifs extensionnels. Enfin, je présente un corpus de test pour l'inférence basée sur les adjectifs que j'ai développée comme ressource pour l'évaluation de systèmes de traitement automatique de l'inférence de la langue naturelle.

Acknowledgments

First of all, I would like to thank my supervisors Claire Gardent and Manfred Pinkal without whom this thesis would not have been possible. They gave me the opportunity to work within an international very stimulating environment. The research work of this thesis in fact has taken place within the *cotutelle* program between the University of Nancy and the University of Saarland.

I warmly thank Claire Gardent for her guidance, patience and invaluable support throughout this thesis. I have benefited from many interesting discussions and from her comments on earlier drafts.

I would like to express my thanks to Manfred Pinkal whose enthusiasm for computational linguistics influenced my views and who gave me the possibility to teach at the University of Saarland which was a great experience for me.

I also thank Petra Dünges for her encouragement in the last months of writing and Stefan Gremm for his helpful collaboration with the test suite.

Finally, I thank my family and all the friends who supported me during these years and especially Massimo, Marco and Dario who gave me their loving support and understanding which helped me writing this thesis.

Table of Contents

Ausführliche Zusammenfassung

Einen Text zu verstehen ist eine der wichtigsten Aufgaben der computergestützten Sprachverarbeitung. Um dieses Ziel zu erreichen, müssen Systeme entwickelt werden, die in der Lage sind eine Bedeutungsrepräsentation für jeden Text zu erstellen und logische Schlussfolgerungen über die Bedeutung eines Textes zu ziehen.
Vor einiger Zeit hat die PASCAL Recognising Textual Entailment Challenge RTE (Dagan et al., 2006) die Aufmerksamkeit der Forscher auf sich gelenkt, und zwar, indem sie die Erkennung von Entailment, eine der wichtigsten Aufgaben beim Ziehen logischer Schlussfolgerungen aus Texten, in den Mittelpunkt ihrer Forschung stellt. Diese Aufgabe besteht darin zu beurteilen, ob angesichts zweier vorliegender Texte T (Text) und H (genannt die Hypothese), der Text T die Hypothese H logisch beinhaltet.
So kann beispielsweise in (1) die Bedeutung des Textes der Hypothese (H) *Joko Ono is John Lennon's widow* (Joko Ono ist John Lennons Witwe) aus der in Text T vorhandenen Information *her late husband* (ihr verstorbener Mann) abgeleitet werden, während in (2) keine Entailment Relation zwischen T und H besteht, da sich die Information *Jupiter has four moons* (Jupiter hat vier Monde) nicht aus *four of Jupiter's tiniest moons* (vier der kleinsten Monde Jupiters) ableiten läßt.

(1) T: *Joko Ono unveiled a bronze statue of her late husband, John Lennon, to complete the official renaming of England's Liverpool Airport as Liverpool John Lennon Airport.*
*(*Zum Abschluß der offiziellen Umbenennung des englischen Liverpool Flughafens in Liverpool John Lennon Flughafen, enthüllte Joko Ono eine Bronzestatue ihres verstorbenen Mannes*)*

H: *Joko Ono is John Lennon's widow.*
*(*Joko Ono ist John Lennons Witwe*)*

(2) T: *They are made from the dust of four of Jupiter's tiniest moons.*
*(*Sie bestehen aus dem Staub von vier der kleinsten Monde Jupiters*)*

H: *Jupiter has four moons.*
*(*Jupiter hat vier Monde*)*

Die Erkennung von Entailment hat ein breites Spektrum von Anwendungsmöglichkeiten. Sie kann dazu dienen festzustellen, ob ein Textfragment eine Frage beantwortet (bei Question Answering Anwendungen), ob eine Abfrage in einem relevanten Dokument enthalten ist (bei Information Retrieval Anwendungen), oder ob ein Textfragment einen spezifischen Textnugget beinhaltet (bei Information Extraction Anwendungen), etc. Da bei diese Anwendungen größtenteils mit beliebigen Texten gearbeitet wird, müssen die

Systeme die mit Entailment Erkennung arbeiten, robust sein. Daher basieren die meisten der bisher entwickelten Systeme auf statistischen Methoden (z.B. stochastisches parsen, Wortabstände oder Wortüberlappungen zur Feststellung semantischer Übereinstimmungen) und nur wenige erlauben eine Integration lexikalischer und kompositioneller Semantik. Jedoch zeigte eines der RTE-1 Forscherteams (Venderwende et al., 2005), dass ungefähr 50% der Entailment Fälle aus dem RTE-1 Datensatz von einem System adäquat bearbeitet werden könnten, wenn dieses semantische Entailment Relationen auf syntaktischer (z.B syntaktische Alternationen) oder lexikalisch semantischer Grundlage (z.B. Synonymie, Antonymie) erfassen würde. Die Tatsache, dass sich die Genauigkeit, bei einer Bezugslinie von 50%, zwischen 50 und 60 Prozent bewegte, weist darauf hin, dass eine verbesserte Integration von Syntax sowie kompositioneller und lexikalischer Semantik die Genauigkeit der Entailmenterkennung verbessern würde. Außerdem erzielten bei der RTE-2 Challenge überraschenderweise zwei Systeme die besten Ergebnisse, die sich auf eine Integration von deep und shallow Techniken (Hickl et al., 2006) sowie auf logische Inferenzen, die auf handschriftlichen Inferenzregeln und Hintergrundwissen (Tatu et al., 2006) basieren, stützen. So wurde zum ersten Mal bewiesen, dass Systeme die mit Tiefenanalyse arbeiten bessere Leistungen bringen als diejenigen, die auf Grundlage statistischer Methoden, funktionieren.

Jedoch wird die Entwicklung dieser und anderer hochwertiger NLP Systeme durch zwei eng verwandte Faktoren erschwert. Zum einen besteht ein Mangel an geeigneten Ressourcen für die Verarbeitung beliebiger Texte, die eine Integration von linguistischem Wissen und Weltwissen bereitstellen. Diese sind notwendig, um logische Schlussfolgerungen und Inferenzaufgaben zu unterstützen. Die Erstellung solcher lexikalischer Ressourcen ist in der Tat eine sehr zeitaufwendige Angelegenheit. Zum anderen macht es das mangelnde Verständnis der Phänomene, die Entailment zugrunde liegen, oft schwierig zu entscheiden, welche Art Wissen notwendig ist, d.h. ob linguistisches Wissen oder ein allgemeines Weltwissen notwendig sind, um mit Inferenzen umgehen zu können. Auch wurde der Begriff des Textual Entailment bisher nicht klar definiert. Dies zeigt sich in der Schwierigkeit, Datensätze für die Evaluierung von Systemen zu entwickeln, die mit Textual Entailment arbeiten. In den bisher veröffentlichten Testdatensätzen (z.B Paraphrase Research Corpus Datensätzen (C. et al., 2004; Dolan et al., 2004), RTE1 (Dagan et al., 2006), RTE2 Bar-Haim et al. (2006) und der Text Retrieval Conference (TREC, 2007) basierte die Beurteilung von Textual Entailment nicht auf semantischer Äquivalenz, sondern auf einem eher informellen Begriff der Semi-Äqivalenz und allgemeinem menschlichen Verständnis, was oft zu Unstimmigkeiten und, wie einige Autoren bemerkten, (Zaenan et al., 2005) sogar zu Fehlern führte. Ich vertrete die Auffassung, dass eine hochwertige computergestützte Bearbeitung von Entailment (d.h. von Entailment Problemen, deren Lösung ausschließlich auf linguistischen Informationen beruht, die dem Text direkt entnommen werden können und nicht auf Weltwissen) unabhängig davon, ob auf statistischer oder symbolischer Grundlage, eine genaue Definition des Begriffes Textual Entailment, entsprechend der des logischen Entailment oder der semantischen Äquivalenz, voraussetzen. Dies wiederum setzt ein tieferes Verständnis der für die (textuellen) Inferenz verantwortlichen Phänomene voraus und eine Lexikontheorie, welche die Erstellung lexikalischer Ressourcen, die das zur Behandlung von Inferenz erforderliche Wissen liefern, unterstützt. Wie einige Autoren feststellten, sind diejenigen Ansätze zur Lexikontheorie geeigneter zur Behandlung von Textual Entailment und NLP im Allgemeinen, die auf der Idee basieren, dass: *die Definition eines Mapping zwischen Wissen und seinem linguistischen Ausdruck erleichtert wird, wenn es möglich ist alle besonderen Beispiele von Tatsachen, Zuständen, Situationen, etc. zu klassifizieren, die in Form einer Gruppe allgemeiner Objekte und spezifizierten Typen von Relationen auftreten und die sich hinsichtlich ihrer möglichen linguistischen Realisierungen, systematisch verhalten* (Bateman, 1990). Tatsächlich haben sich lexikalisch-semantische Verbklassifikationen in NLP Anwendungen als nützlich erwiesen (Dorr, 1997; Kipper et al., 2000; Klavans and Kan, 1998; Stede, 1998), da sie Generalisierungen erlauben und so helfen Redundanzen im Lexikon zu vermeiden. Darüberhinaus unterstützen sie die automatische Extraktion großvolumiger lexikalischer Ressourcen aus Corpora (Brew and im Walde, 2002; Briscoe and Carroll, 1997; Merlo and Stevenson, 2001).

Eine der wichtigsten Arbeiten auf diesem Gebiet ist Levins (1993) Klassifizierung englischer Verben. Basierend auf der Idee, dass syntaktisches und semantisches Verhalten eng miteinander verbunden sind, definiert Levin Verbklassen auf der Basis ihrer syntaktischen Alternationen. Folglich teilen alle Elemente in einer Verbklasse dieselbe Menge Alternationen, d.h. paraphrastischer Konstruktionen, während Verben die verschiedenen Klassen angehören, sich bezüglich der für sie möglichen paraphrastischen Konstruktionen (Alternationen) unterscheiden. So gehören beispielsweise die Verben *meet* (treffen) und *embrace* (umarmen) verschiedenen Klassen an, da bei *meet* ein Weglassen der Präposition *with* (3) möglich ist, während dies bei *embrace* (4) nicht der Fall ist.

*(3) a. Anne met with Cathy (*Anne traf sich mit Cathy*)*

*b. Anne met Cathy (*Anne traf Cathy*)*

*(4) a. *Brenda embraced with Molly (*Brenda und Molly umarmten sich*)*

*b. Brenda embraced Molly (*Brenda umarmte Molly*)*

Diese Idee ist sehr vielversprechend für die Erkennung von Entailment und speziell für die Paraphrasenerkennung, da Alternationen durch dieselbe semantische Repräsentation ausgedrückt werden können und darüberhinaus eine automatische Extraktion paraphrastischer Verbmuster ermöglicht wird, indem ausschließlich im Lexikon vorhandene Informationen verwendet werden.

Daher untersuche ich in der vorliegenden Dissertation linguistisch-basierte Ansätze zur Erkennung von Textual Entailment und Paraphrasen, die auf einer feinkörnigen semantischen Analyse beruhen. Insbesondere untersuche ich, inwiefern lexikalische Ressourcen, die auf semantischen Klassifikationen von grammatischen Kategorien beruhen, wie die Levins, erweitert werden können, so dass sie das zur Behandlung von Inferenz notwendige Wissen liefern. Dabei setze ich den Schwerpunkt meiner Arbeit auf verbbasierte textuelle Äquivalenz und adjektivbasiertes Entailment.

Verbparaphrasenerkennung

Ein auffälliges Charakteristikum natürlicher Sprache ist, dass sie Paraphrasen gestattet, d.h. sie gestattet verschiedene Verbalisierungen desselben Inhalts. Folglich teilen sie, obwohl die verschiedenen Verbalisierungen in (5) unterschiedliche pragmatische oder kommunikative Werte haben (z.B. hinsichtlich Topikalisierung, Präsupposition oder Fokus-Hintergrund Partitionierung), einen gemeinsamen semantischen Kern(inhalt): der Inhalt der durch die traditionelle montaguesche kompositionelle Semantik beschrieben wird.

(5) a. *The cruise is expensive.*
(Die Kreuzfahrt ist teuer)

b. *The cost of the cruise is high.*
(Die Kreuzfahrt hat hohen Preis)

c. *The cruise has a high cost.*
(Die Kreuzfahrt hat hohen Preis)

In natürlicher Sprache gibt es zahlreiche Möglichkeiten paraphrastische Strukturen zu erzeugen. Reformulierungen derselben Bedeutung können unterschiedliche Quellen haben, abhängig davon, welche Art von Wissen die Quelle der textuellen Äquivalenz ist. Ich unterscheide zwischen lexiko-strukturellen Paraphrasen, bei denen die Ursache der textuellen Äquivalenz entweder lexikalisches Wissen oder Wissen über die beteiligten syntaktischen Relationen im Paraphrasenpaar ist und extra-linguistischen Paraphrasen, bei denen der Ursprung der Äquivalenz in Weltwissen oder im Kontext der Äußerung liegt. So sind beispielsweise die Sätze in (5) lexiko-strukturelle Paraphrasen, da die Relationen zwischen den beiden Sätzen durch verschiedenen syntaktische Realisierungen erklärt werden können. In Beispiel (6) jedoch, wird der referentielle Ausdruck durch seinen Referenten ersetzt.

(6) a. *I wrote this letter two months ago.*
(Ich schrieb diesen Brief vor zwei Monaten)

b. *I wrote this letter in August.*
(Ich schrieb diesen Brief im August)

In diesem Fall reicht die lexikalische Semantik allein nicht aus, um zu folgern, dass es sich bei den beiden Sätzen um Paraphrasen handelt. Hier sind Wissen über den Kontext und die genaue Bedeutung des referentiellen Ausdrucks notwendig. Ähnliches gilt für die Sätze in (7), bei denen es sich eindeutig um Paraphrasen handelt, der Ursprung der Äquivalenz jedoch abermals weder lexikalisch noch strukturell begründet ist. In diesem Fall dient Weltwissen zur Feststellung der semantischen Äquivalenz, d.h. das Wissen, dass neugeborene Kaninchen noch keine Zähne haben.

(7) a. *I found a newborn rabbit in the garden.*
*(*Ich fand ein neugeborenes Kaninchen im Garten*)*

b. *I found a newborn rabbit with no teeth in the garden.*
*(*Ich fand ein neugeborenes Kaninchen ohne Zähne im Garten*)*

Die lexiko-strukturelle textuelle Äqivalenz steht im Mittelpunkt der vorliegenden Arbeit. Linguisten stellten bereits vor langer Zeit fest, dass Paraphrasen in natürlicher Sprache weit verbreitet sind und versuchten sie zu charakterisieren. So erfassen beispielsweise Chomskys Transformationen die Beziehung zwischen einer Kernbedeutung (in Chomskys Worten Tiefenstruktur) und verschiedenen syntaktischen Realisierungen (z.B. zwischen der Passivform und der Aktivform desselben Satzes), während Melcuk (1988) sechzig paraphrastische Regeln aufstellt, um die paraphrastischen Relationen zwischen Sätzen beschreiben zu können. In letzter Zeit hat die Arbeit auf dem Gebiet der Informationsextraktion (IE) und Question Answering (QA) ein neues Forschungsinteresse hinsichtlich der Paraphrasen geweckt, da IE und QA Systeme in der Lage sein müssen, verschiedene Verbalisierungen desselben Inhalts zu erkennen. Aufgrund der open domain Corpora, mit denen diese Systeme klarkommen müssen, sind Coverage und Robustheit die Kernprobleme, und ein großer Teil der Arbeit auf dem Gebiet der Paraphrasen basiert auf automatischen Lerntechniken. Der hier vorgestellte Ansatz zur Paraphrasenerkennung ist linguistisch basiert und zielt darauf ab, ein System zu erstellen, das in der Lage ist zwei unterschiedlichen paraphrastischen Konstruktionen dieselbe semantische Repräsentation zuzuordnen. Dieser Ansatz stellt die Erkennung von Verbparaphrasen in den Mittelpunkt und baut auf Levins (1993) Klassifizierung englischer Verben auf.

Inspiriert durch die Idee, dass ähnliche syntaktische Eigenschaften der Schlüssel zu ähnlichem semantischem Verhalten sind, konstruiert Levin semantische Klassen englischer Verben, indem sie den Schwerpunkt auf syntaktische Alternationen setzt, d.h. unter Berücksichtigung aller unterschiedlichen syntaktischen Muster, die für Verben möglich sind, die mit demselben semantischen Inhalt in Zusammenhang stehen.

(8) a. *The key opens the safe* ↔ *The safe opens with the key*
*(*Der Safe lässt sich mit dem Schlüssel öffnen*)*

b. *I give books to John* ↔ *I give John books*
*(*Ich gebe John Bücher*)*

Folglich haben Verben die derselben Verbklasse angehören dieselben semantischen Eigenschaften. In der vorliegenden Dissertation stelle ich die Behauptung auf, dass die von Levin vorgeschlagene spezielle Organisation der Bedeutung von Verben für die automatische Beschaffung und Erkennung paraphrastischer Verbmuster geeignet ist. Dies gilt insbesondere für die Tatsache, dass jede Verbklasse und folglich jedes Verb mit einer Reihe von Alternationen in Zusammenhang stehen, d.h. paraphrastische Muster die für die Elemente einer Klasse möglich sind. Daher zeige ich, wie die paraphrastischen Muster, die sich aus Levins Klassifikation ergeben, durch linguistische Informationen über Synonyme, Troponyme, etc. aus WordNet (Fellbaum, 1998) erweitert und für die automatische Erfassung paraphrastischer Verbmuster genutzt werden können. Außerdem zeige ich wie diese Informationen mit einem robusten Parser wie XIP (Xerox Incremental Parser (Ait-Mokhtar et al., 2002)) verbunden werden können und wie sie genutzt werden können um, auf Alternation beruhende Paraphrasen zu erkennen. Der Parser wurde mit Hilfe einer Test Suite, die aus paraphrastischen Satzpaaren besteht evaluiert und das mit einem ermutigenden Ergebnis, d.h. die textuelle Äquivalenz wurde in 96% der Fälle erkannt. Der hier vorgestellte symbolische Ansatz zur Paraphrasenbehandlung bedeutet einen ersten Schritt in Richtung einer robusten symbolischen Handhabung von Paraphrasen. Tatsächlich ermöglicht dieser Ansatz eine Extraktion aus beliebigen Texten und die Erkennung paraphrastischer Verbmuster, indem aus großvolumigen lexikalischen Ressourcen wie VerbNet und WordNet extrahierte linguistische Informationen mit einem robusten Parser wie XIP verbunden werden.

Adjektivbasiertes Entailment

Levin (1993) stützt ihre Klassifikation englischer Verben auf eine Gruppe von 79 Alternationen, welche die paraphrastischen Muster, in denen die Verben vorkommen können, beschreiben. Basierend auf der Beobachtung, dass Paraphrasen eine Untergruppe einer allgemeineren Gruppe von Inferenzmustern sind, schlage ich eine wichtige Erweiterung für Levins System vor, und zwar behaupte ich, dass linguistische Merkmale besser beschrieben werden können, wenn man die Gruppe der von ihnen gezeigten Inferenzmuster berücksichtigt. Außerdem beruhen diese Muster, im Gegensatz zu Levins Alternationen; auf linguistischen Prinzipien, da jeder Inferenztypus durch seine ursächliche linguistische Eigenschaft, begründet ist. Dann habe ich diese Idee auf die Klassifizierung englischer Adjektive anwendet. In der lexikalischen Semantik haben Adjektive, obgleich sie eine sehr interessante grammatische Kategorie repräsentieren, bisher nicht die gleiche Aufmerksamkeit erfahren wie Verben und Nomen. Tatsächlich zeigen Adjektive keinen homogenen Satz von Eigenschaften.
Syntaktisch können sie in verschiedenen Alternationen auftreten, z.B: in attributiven und prädikativen Konstruktionen (z.B. *red*, *rectangular*) oder nur in einer der beiden Konstruktionen (z.B. *asleep*, *gastronomical*). Sie können obligatorische (z.B. *reluctant*) oder fakultative (z.B. *sad*) Komplemente haben oder gar keine Komplemente zulassen (z.B. *rectangular*). Einige Adjektive sind steigerbar (z.B. *big*, *quick*), andere sind es nicht (z.B. *finite, carnivorous*). Semantisch scheinen Adjektive ein unterschiedliches Verhalten zu zeigen, einige Adjektive (z.B. *red*, *big*) sind extensional, d.h. sie bezeichnen Eigenschaften erster Ordnung, andere (z.B. *alleged, former*) sind intensional und zeigen Funktionen, die Eigenschaften von Eigenschaften abbilden, d.h. Eigenschaften zweiter Ordnung. Es gibt viele theoretische Ansätze zur Klassifizierung von Adjektiven. Bei einigen steht die Analyse ihrer syntaktischen Eigenschaften im Mittelpunkt (siehe Huddleston, 1984; Quirk et al., 1985; Vendler, 1963, 1968), andere setzen den Schwerpunkt auf eine Analyse der modelltheoretischen Eigenschaften (siehe Kamp, 1975; Keenan and Faltz, 1985; Keenan, 1987; Chierchia and Connell-Ginet, 1990; Kamp and Partee, 1995) und wieder andere konzentrieren sich auf ontologische Unterscheidungen (siehe

Dixon, 1982, 1991; Frawley, 1992; Aarts, 1976). Die wahrscheinlich interessanteste Klassifikation von Adjektiven hinsichtlich der Erkennung von Entailment ist die von Kamp und Partee (Kamp, 1975; Kamp and Partee, 1995). Sie unterteilen englische Adjektive in vier semantische Klassen, indem sie die Entailmentmuster modelltheoretisch-semantischen Ursprungs, in denen die Adjektive auftreten, in ihre Betrachtung miteinbeziehen. Folgt man dieser Klassifikation, so ist ein Adjektiv wie *red* intersective, da die [A N]$_{AP}$ Phrase die Verbindung der Eigenschaften bedingt, die durch das Nomen N und das Adjektiv A ausgedrückt werden. Ein Adjektiv wie *big* ist subsective, da die [A N]$_{AP}$ Phrase nur die durch das Nomen N ausgedrückte Eigenschaft bedingt. Ein Adjektiv wie *fake* ist privative, da die Phrase die Negation der durch N ausgedrückten Eigenschaft bedingt. Und ein Adjektiv wie *alleged* ist plain nonsubsective, da sich nichts aus der [A N]$_{AP}$ ableiten läßt.
Jedoch gelingt es keiner dieser Klassifikationen, die Komplexität des Verhaltens von Adjektiven vollständig zu beschreiben. Das Haupthindernis ist, dass sie das Zusammenspiel zwischen den verschiedenen Stufen linguistischer Beschreibung nicht berücksichtigen, z.B. Syntax, Semantik, lexikalische Semantik und derivationelle Morphologie, die alle wichtige Faktoren bei der Bestimmung möglicher Inferenzen sind. Wie die Beispiele (9) und (10) zeigen, weisen die Adjektive *dry* (trocken) und *rectangular* (rechteckig) unterschiedliche Inferenzmuster hinsichtlich der lexikalisch-semantischen Eigenschaft Antonymie auf, obwohl beide (modelltheoretisch) semantisch als intersective klassifiziert werden können.

(9) a. *The dishcloth is dry → The dishcloth is not wet*
*(*Das Geschirrtuch ist trocken → Das Geschirrtuch ist nicht nass*)*

b. *The dishcloth is not dry → The dishcloth is wet*
*(*Das Geschirrtuch ist nicht trocken → Das Geschirrtuch ist nass*)*

(10) a. *This table is rectangular → This table is not round*
*(*Dieser Tisch ist rechteckig → Dieser Tisch ist nicht rund*)*

b. *This table is not rectangular ↛ This table is round*
*(*Dieser Tisch ist nicht rechteckig ↛ Dieser Tisch ist rund*)*

Das Adjektiv *dry* geht eine binäre antonymische Relation mit seinem Antonym *wet* ein. Das bedeutet, die Entailment-Relation zwischen dem Adjektiv und der Negation seines Antonyms gilt in beide Richtungen, z.B. $A \models \neg Anto(A) \wedge \neg A \models Anto(A)$. Im Gegensatz dazu, hat das Adjektiv *rectangular* kein eindeutiges Antonym und geht multiple Relationen mit der Gruppe seiner Antonyme ein. In diesem Fall gilt die Relation zwischen dem Adjektiv und der Negation eines seiner Antonyme nur in einer Richtung, z.B. $A \models \neg Anto_1(A) \wedge \neg A \not\models Anto_1(A)$.
Dies weist darauf hin, dass die von Kamp und Partee vorgeschlagenen vier Adjektivklassen in Unterklassen aufgespaltet werden sollten und dass auch andere Eigenschaften, neben den rein modelltheoretischen, berücksichtigt werden sollten. Deshalb habe ich in

der vorliegenden Dissertation zuerst eine Gruppe von Charakteristika bestimmt, die für die Beschreibung des Verhaltens von Adjektiven unabdingbar sind. Diese Charakteristika wurden erhalten durch eine Kombination von Auswahlkriterien, wie sie in traditionellen Adjektivklassifikationen verwendet werden (Huddleston, 1984; Quirk et al., 1985; Vendler, 1963, 1968; Kamp, 1975; Keenan and Faltz, 1985; Keenan, 1987; Chierchia and Connell-Ginet, 1990; Kamp and Partee, 1995) und lexikalisch semantischen sowie morphoderivationellen Kriterien. Dann assoziiere ich jede Eigenschaft mit den Inferenzmustern, die sie generiert, und ich klassifiziere eine Auswahl von 500 Adjektiven, indem ich die Gruppe der Inferenzmuster, in denen die Elemente der Gruppe auftreten können, als Auswahlkriterien verwende. Folglich wird eine semantische Klasse durch die Gruppe der Inferenzmuster, in der ihre Elemente auftreten können, definiert. Das Ergebnis ist eine aus 42 Klassen bestehende, feingekörnige semantische Klassifikation englischer Adjektive, die es ermöglicht, ihr inferentielles Verhalten besser vorherzusagen. Weiterhin entwickle ich, basierend auf Hobbs (1985) ontologisch gemischten Ansatz, eine kompositionelle semantische Systematik um Adjektiven eine semantische Repräsentation zuzuweisen und ermögliche dadurch eine Inferenz erster Ordnung für alle Typen von Adjektiven, einschließlich der extensionalen Adjektive. Wie Hobbs (1985) und andere argumentierten, muss die semantische Repräsentation natürlicher Sprache nicht höherer Ordnung sein, da das Problem durch die ontologische Promiskuität (ontological promiscuity) gelöst werden kann. Das bedeutet, indem man alle Objekte, die man durch ein Prädikat beschreiben kann, verdinglicht, ist es möglich ein semantisches Repräsentationsschema erster Ordnung für natürliche Sprache zu erhalten. Diese Beobachtung ist äußerst wichtig für computergestützte Sprachsysteme, und zwar aus zwei Gründen. Erstens ist eine Logik, die über eine Logik erster Ordnung hinausgeht, nicht entscheidbar. Der zweite und wichtigere Grund ist, dass es keine gängigen automatischen Beweissysteme gibt, die man einsetzen könnte, um über Bedeutung von Formeln höherer Ordnung logisch zu schlussfolgern. Die semantische Repräsentation, die den Adjektiven zugewiesen wird, berücksichtigt auch ihre morphoderivationellen und lexikalisch semantischen Eigenschaften. Zur Evaluierung der vorgeschlagenen Klassifikation, erstelle ich zuerst eine Test Suite, bestehend aus Satzpaaren, die die relevanten Inferenzmuster von Adjektiven veranschaulichen. Dann benutze ich das BB1 (Blackburn and Bos, 2005) semantische Rechensystem, um, erstens, eine semantische Repräsentation mit den Sätzen in der Test Suite zu assoziieren und, zweitens, um zu testen, ob die Elemente jedes Satzpaares in einer Entailment Relation stehen oder nicht. Die Ergebnisse zeigen, dass die vorgeschlagene Semantik die Inferenzmuster, die für die verschiedenen in der vorliegenden Arbeit identifizierten Unterklassen gelten, korrekt prognostiziert.

Schlussfolgerungen

In der vorliegenden Dissertation habe ich untersucht wie lexikalische Ressourcen, die auf der Gliederung lexikalischen Wissens in Klassen mit gemeinsamen Eigenschaften (lexikalische, semantische etc,) basieren, die computergestützte Verarbeitung natürlicher

Sprache und insbesondere die symbolische Erkennung von Entailment unterstützen. Basierend auf Levins (1993) Klassifikation englischer Verben, wurde zuerst ein robuster, für die Verarbeitung beliebiger Texte geeigneter Ansatz zur Paraphrasenerkennung vorgestellt. Dann habe ich aufgezeigt, dass man durch eine Erweiterung von Levins Systematik zur Behandlung allgemeiner Inferenzmuster, eine Klassifikation von englischen Adjektiven erhält, die verglichen mit früheren Ansätzen, eine feinkörnige semantische Charakterisierung ihrer inferentiellen Eigenschaften gestattet und so die Basis für die computergestützte Behandlung von Inferenz bei Adjektiven bildet. Der Hauptbeitrag der vorliegenden Arbeit besteht in der detaillierten Analyse der Wechselbeziehungen zwischen derivationeller Morphologie und lexikalischer sowie kompositioneller Semantik und deren Auswirkungen auf die Entailmentmuster, welche Sätze die Adjektive und deren verwandte Nomen und Verben enthalten, erlauben. Der hier vorgestellte Ansatz schafft die Grundlage für eine computergestützte Behandlung adjektivbasierter Inferenz, indem hierbei eine feinkörnige Charakterisierung der verschiedenen Inferenzmuster, die für Adjektive gestattet sind, bereitgestellt wird. Ich glaube, dass die in dieser Dissertation vorgestellte Methodologie für die Klassifizierung von Adjektiven, d.h. die Definition semantischer Klassen durch die Gruppe von Inferenzmustern, die ihre Elemente erlauben, vielversprechend ist für die Schaffung allgemeiner, linguistisch basierter Wissensontologien (Bateman, 1990), die eine semantisch syntaktische Schnittstelle für Domain orientierte Wissensontologien liefern. Ein anderes beachtliches Ergebnis ist die Test Suite die ich entwickelt habe und die als Ressource für NPL Anwendungen, die Inferenzen (insbesondere Inferenzen bei Adjektiven) behandeln, genutzt werden kann. Im Gegensatz zu anderen Corpora, die zu diesem Zweck entwickelt wurden, wie etwa die Datensätze des Paraphrase Research Corpus (C. et al., 2004; Dolan et al., 2004), der RTE1 (Dagan et al., 2006), RTE2 (Bar-Haim et al., 2006) und der Text Retrieval Conference TREC (2007), hat diese Test Suite einen klar definierten Anwendungsbereich und enthält Annotationen für Inferenzmuster. Dadurch wird eine bessere Kontrolle der analysierten Daten ermöglicht. Durch die Konstruktion dieser Test Suite beabsichtige ich, den Weg für die Schaffung von Ressourcen zu ebnen, die einen tieferen Einblick in die für Inferenz verantwortlichen Phänomene ermöglichen.

Résumé détaillé

La compréhension d'un texte constitue l'un des objectifs les plus importants du traitement automatique des langues. Pour atteindre cet objectif, des systèmes capables d'élaborer une représentation de la signification de n'importe quel texte et de raisonner sur la signification d'un texte doivent être développés. Récemment, le défi RTE (Recognising Textual Entailment, reconnaissance d'implications textuelles) PASCAL (Dagan et al., 2006) a suscité une attention particulière au sein de la communauté des chercheurs en se concentrant sur l'une des principales tâches d'inférence impliquées dans le raisonnement textuel, à savoir la tâche de reconnaissance d'implications. Cette tâche consiste à juger si, selon deux textes T (texte) et H (appelé hypothèse), le texte T implique l'hypothèse H. Dans l'exemple (1), la signification évoquée par le texte de l'hypothèse (H), c'est-à-dire que Joko Ono est la veuve de John Lennon, peut être déduite de l'information son regretté mari, John Lennon contenue dans le texte T, alors que dans l'exemple (2), l'implication entre T et H n'est pas valable car l'information Jupiter a quatre lunes ne découle pas de quatre des plus petites lunes de Jupiter.

(1) T: *Joko Ono unveiled a bronze statue of her late husband, John Lennon, to complete the official renaming of England's Liverpool Airport as Liverpool John Lennon Airport.*
*(*À la fin de la cérémonie officielle organisée pour rebaptiser l'aéroport Liverpool Airport, en Angleterre, en Liverpool John Lennon Airport, Joko Ono a dévoilé une statue en bronze de son défunt mari, John Lennon.*)*

H: *Joko Ono is John Lennon's widow.*
*(*Joko Ono est la veuve de John Lennon*)*

(2) T: *They are made from the dust of four of Jupiter's tiniest moons.*
*(*Elles sont faites de poussière de quatre des plus petites lunes de Jupiter*)*

H: *Jupiter has four moons.*
*(*Jupiter a quatre lunes*)*

La reconnaissance d'implications comporte une large gamme d'applications. Elle peut être utilisée pour déterminer si un fragment de texte répond à une question (dans une application de questions-réponses), si une question est impliquée par un document pertinent (dans une extraction d'information), si un fragment de texte implique une partie spécifique d'informations (dans une extraction d'information), etc. Dans la mesure où la plupart de ces applications se concentrent sur un texte réel, les systèmes chargés de la reconnaissance d'implications doivent être robustes. Par conséquent, la plupart des systèmes développés jusqu'à présent sont basés sur des méthodes statistiques (par ex., l'analyse syntaxique stochastique et la distance lexicale ou le chevauchement lexical pour une similarité sémantique) et quelques-uns d'entre eux permettent une intégration de sémantique lexicale et compositionnelle. Toutefois, l'une des équipes RTE-1 participantes (Venderwende et al., 2005) a montré que environ 50% des cas d'implications présents dans le jeu de données RTE-1 pourraient être traités correctement par un système qui couvrirait de manière adéquate les implications sémantiques basées sur la syntaxe (par ex. les alternances syntactiques) ou basées sur la sémantique lexicale (par ex. la synonymie, l'antonymie). Le fait que les exactitudes moyennes du système s'élèvent entre 50 et 60 % avec une ligne de base de 50% insinue qu'une meilleure intégration de la syntaxe et de la sémantique compositionnelle et lexicale pourrait améliorer l'exactitude de la reconnaissance d'implications. De plus, les meilleurs résultats du défi RTE-2 ont étrangement été obtenus par deux systèmes basés sur l'intégration de techniques deep et shallow (Hickl et al., 2006) et sur l'inférence logique à l'aide de règles d'inférence manuscrites et de connaissances de fond (Tatu et al., 2006). Par conséquent, ceci démontre pour la première fois que les systèmes qui traitent l'analyse approfondie fonctionnent mieux que les systèmes basés sur les statistiques. Cependant, le développement de tels systèmes et de systèmes de traitement de la langue naturelle (NLP) de haute qualité en général est rendu difficile par deux facteurs proches l'un de l'autre. Premièrement, il existe un manque de ressources à large couverture appropriées et offrant une intégration de connaissances linguistiques et de base nécessaires pour permettre les tâches de résonnement et d'inférence. L'élaboration de telles ressources lexicales est en effet une tâche qui requiert énormément de temps. Deuxièmement, le manque de compréhension approfondie des phénomènes qui sont à l'origine de l'implication rend souvent difficile le fait de décider quel type de connaissance, c'est-à-dire la connaissance linguistique ou plus généralement la connaissance du monde, est nécessaire pour traiter une inférence. De même, la notion d'implication textuelle n'a jusqu'à présent pas été clairement définie. Ceci se reflète dans la difficulté de produire des corpus de test d'évaluation pour les systèmes qui traitent l'implication textuelle. Dans les jeux de données d'essai (par ex. les jeux de données Paraphrase Research Corpus (C. et al., 2004; Dolan et al., 2004), RTE1 (Dagan et al., 2006), RTE2 (Bar-Haim et al., 2006) et les données Text Retrieval Conference TREC (2007)) qui ont été publiés jusqu'à présent, l'évaluation de l'implication n'est pas basée sur l'équivalence sémantique mais sur une notion plus « informelle » de semi-équivalence, de sens commun, de compréhension humaine commune, ce qui entraîne souvent des divergences et, comme certains auteurs l'ont remarqué (Zaenan et al., 2005), même des erreurs. Je pense qu'un traitement automatique de haute qualité de l'implication textuelle (c'est-à-dire de problèmes d'implication pouvant être résolus uniquement sur la base des informations linguistiques qui peuvent être extraites directement du texte et non pas sur la base de la connaissance du monde), qu'il soit statistique ou symbolique, requiert la définition précise de la notion d'implication textuelle conformément à celle de l'implication logique ou de l'équivalence sémantique. Ceci présuppose une compréhension plus profonde des phénomènes linguistiques responsables de l'inférence (textuelle) et une théorie du lexique qui permet la création de ressources lexicales fournissant la connaissance nécessaire pour traiter l'inférence. Comme certains

auteurs l'ont constaté, les approches de la théorie du lexique qui sont plus intéressantes pour l'implication textuelle et le traitement de la langue naturelle (NLP) en général sont celles qui sont basées sur l'idée que *la définition d'un mappage entre la connaissance et son expression linguistique est facilitée s'il est possible de classifier tous les cas particuliers de faits, d'états, de situations, etc, qui se produisent sous forme d'un ensemble d'objets généraux et de relations de types spécifiés qui se comportent systématiquement quant à leurs réalisations linguistiques possibles* (Bateman, 1990).
En effet, les classifications sémantiques lexicales de verbes ont été démontrées comme étant utiles dans le traitement de la langue naturelle (NLP) (Dorr, 1997; Kipper et al., 2000; Klavans and Kan, 1998; Stede, 1998) car elles permettent des généralisations et par conséquent évitent la redondance dans le lexique. De plus, ces classifications permettent d'extraire automatiquement de corpus des ressources lexicales à large couverture (Brew and imWalde, 2002; Briscoe and Carroll, 1997; Merlo and Stevenson, 2001).

L'un des plus importants travaux effectués dans ce domaine est la classification des verbes anglais par Levin (1993). Selon l'idée que les comportements syntactiques et sémantiques sont étroitement liés, Levin définit un groupe de classes de verbes sur la base de leurs alternances. Par conséquent, tous les éléments d'une classe de verbes partagent le même groupe d'alternances, c'est-à-dire de constructions paraphrastiques, alors que les verbes de différentes classes diffèrent dans les constructions paraphrastiques (alternances) possibles pour eux. Ainsi, par exemple, les verbes *meet* (rencontrer) et *embrace* (étreindre) appartiennent à différentes classes, car *meet* permet l'alternance consistant à supprimer la préposition *with* (avec) (3), alors que *embrace* ne le permet pas (4).

(3) a. *Anne met with Cathy*
(Anne a rencontré Cathy)

b. *Anne met Cathy*
(Anne a rencontré Cathy)

(4) a. **Brenda embraced with Molly*
(*Brenda a étreint avec Molly)

b. *Brenda embraced Molly*
(Brenda a étreint Molly)

Cette idée est très prometteuse pour la reconnaissance d'implications, et en particulier pour la reconnaissance de paraphrases, car elle permet d'associer la même représentation sémantique à deux constructions qui s'alternent et d'extraire automatiquement des modèles de verbes paraphrastiques, en se basant uniquement sur l'information disponible dans le lexique. Par conséquent, j'étudie dans cette thèse des approches linguistiques de reconnaissance d'implications textuelles et de paraphrases sur la base d'une sémantique à grain fin. J'étudie en particulier jusqu'où les ressources lexicales basées sur les classifications sémantiques lexicales de catégories grammaticales, telles que celle de Levin, peuvent être étendues afin de fournir la connaissance requise pour traiter une inférence.

Je me concentre sur l'équivalence textuelle basée sur le verbe et sur l'implication basée sur les adjectifs.

Reconnaissance de paraphrases verbales

L'un des aspects frappants de la langue naturelle est qu'elle autorise les paraphrases, c'est-à-dire qu'elle permet différentes verbalisations du même contenu. Par conséquent, bien que les différentes verbalisations de l'exemple (5) peuvent avoir différentes valeurs pragmatiques ou communicatives (par ex. en ce qui concerne la topicalisation, les présuppositions ou la focalisation/partition de base), elles partagent toutes un contenu sémantique essentiel : le contenu représenté par la sémantique compositionnelle montagovienne traditionnelle.

(5) *a.* *The cruise is expensive.*
(La croisière est chère.)

b. *The cost of the cruise is high.*
(Le coû de la croisière est élevé.)

c. *The cruise has a high cost.*
(La croisière a un coû élevé.)

Dans la langue naturelle, il existe de nombreux moyens de produire des structures paraphrastiques. Les reformulations du même sens peuvent avoir différentes sources en fonction du type de connaissance qui est la source de l'équivalence textuelle. Je fais la distinction entre les paraphrases lexico-structurelles dont la source de l'équivalence textuelle est la connaissance lexicale ou la connaissance des relations syntactiques impliquées dans la paire de paraphrases, et les paraphrases extra-linguistiques dont la source de l'équivalence est la connaissance du monde ou le contexte de l'énoncé. Les phrases de l'exemple (5), par exemple, sont des paraphrases lexico-structurelles car la nature des relations qui existent entre les deux phrases peut être expliquée par le biais de différentes réalisations syntactiques. Mais dans l'exemple (6), une expression référentielle est remplacée par son référent.

(6) *a.* *I wrote this letter two months ago.*
(J'ai écrit cette lettre il y deux mois.)

b. *I wrote this letter in August.*
(J'ai écrit cette lettre en aoû)

Dans ce cas, la sémantique lexicale n'est pas suffisante pour en déduire que les deux phrases sont des paraphrases. La connaissance du contexte et de la signification précise de l'expression référentielle est nécessaire. De même, les phrases de l'exemple (7) sont des paraphrases évidentes, mais une fois de plus, la source de l'équivalence sémantique n'est ni lexicale ni structurelle. Dans ce cas, la connaissance du monde est utilisée pour établir l'équivalence sémantique, à savoir la connaissance du fait qu'un lapin qui vient de naître est trop jeune pour avoir des dents.

(7) a. *I found a newborn rabbit in the garden.*
(Dans le jardin, j'ai trouvé un lapin qui vient de naître.)

b. *I found a newborn rabbit with no teeth in the garden.*
(Dans le jardin, j'ai trouvé un lapin sans dents qui vient de naître.)

Dans cette thèse, je me concentre sur l'équivalence textuelle lexico-structurelle. Les linguistes ont depuis longtemps constaté que les paraphrases sont très répandues dans la langue naturelle et ils ont tenté de les caractériser. Ainsi, par exemple, les transformations de Chomsky captent la relation qui existe entre une signification principale (une structure profonde dans les termes de Chomsky) et plusieurs réalisations syntactiques (par exemple, entre les formes passives et actives de la même phrase) alors que Melcuk (1988) présente soixante règles paraphrastiques pour pouvoir décrire les relations paraphrastiques entre les phrases. Plus récemment, le travail sur l'extraction d'information (IE) et la question-réponse (QA) a déclenché un regain d'intérêt dans la recherche sur les paraphrases car les systèmes IE et QA doivent généralement être capables de reconnaître différentes verbalisations du contenu. En raison des vastes outils de langage non contrôlé avec lesquels ces systèmes ont affaire, la large couverture et la robustesse sont des questions fondamentales et la plupart du travail effectué dans ce domaine sur les paraphrases est basé sur des techniques d'apprentissage automatiques. L'approche sur la reconnaissance de paraphrases que je présente dans cette thèse est plutôt différente, elle est basée sur la linguistique et vise à modéliser un cadre capable d'attribuer une seule et même représentation sémantique à deux constructions paraphrastiques différentes. Mon approche se concentre sur la reconnaissance de paraphrases verbales et est construite sur la classification des verbes anglais par Levin (1993).

Inspiré par l'idée que des propriétés syntactiques semblables constituent la clé d'un comportement sémantique semblable, Levin construit des classes sémantiques de verbes anglais en se concentrant sur les alternances syntactiques, c'est-à-dire en prenant en compte tous les modèles syntactiques possibles pour les verbes qui sont associés au même contenu sémantique:

(8) a. *The key opens the safe* ↔ *The safe opens with the key*
(La clé ouvre le coffre-fort ? Le coffre-fort s'ouvre avec la clé)

b. *I give books to John* ↔ *I give John books*
(Je donne des livres à John ? Je donne à John des livres)

Par conséquent, les verbes appartenant à la même classe partagent des propriétés sémantiques semblables. Dans cette thèse, je déclare que l'organisation particulière de la signification verbale proposée par Levin, et en particulier le fait que chaque classe de verbes et par conséquent chaque verbe est associé à un ensemble d'alternances, c'est-à-dire les modèles paraphrastiques possibles pour les membres de la classe, est appropriée pour être utilisée pour l'acquisition et la reconnaissance automatique de modèles de verbes paraphrastiques. Par conséquent, je montre comment les modèles paraphrastiques résultant de la classification de Levin peuvent être étendus par une information linguistique sur les synonymes, troponymes, etc. extraite de WordNet (Fellbaum, 1998), et utilisés pour l'acquisition automatique de modèles de verbes paraphrastiques. De plus, je montre comment cette information peut être reliée à un analyseur robuste tel que XIP (Xerox Incremental Parser (Ait-Mokhtar et al., 2002)) et comment elle peut être utilisée pour reconnaître des paraphrases basées sur l'alternance. L'analyseur a été évalué à l'aide d'un corpus de test de paires de phrases paraphrastiques et a permis d'obtenir un résultat encourageant. En effet, l'équivalence textuelle a été reconnue dans 96% des cas. L'approche symbolique sur la paraphrase que je présente dans cette thèse constitue une première étape vers un traitement symbolique robuste de paraphrases. En effet, l'approche proposée permet une extraction à large couverture et une reconnaissance de modèles de verbes paraphrastiques en intégrant l'information linguistique extraite des ressources lexicales à large couverture existantes telles que VerbNet et WordNet avec un analyseur robuste tel que XIP.

Implication basée sur les adjectifs

Levin (1993) base sa classification des verbes anglais sur un ensemble de 79 alternances décrivant les modèles paraphrastiques auxquels les verbes peuvent participer. Tout en observant le fait que les paraphrases constituent un sous-ensemble de l'ensemble plus général de modèles d'inférence que les éléments linguistiques peuvent présenter, je propose une extension importante du cadre de Levin, c'est-à-dire que je prétends que les aspects linguistiques peuvent être mieux décrits en prenant en compte l'ensemble des modèles d'inférence qu'ils présentent. De plus, contrairement aux alternances de Levin, ces modèles sont réglés par des principes linguistiques car chaque type d'inférence est motivé par la propriété linguistique qui constitue sa source. Puis j'ai appliqué cette idée à la classification sémantique d'adjectifs anglais. En sémantique lexicale, les adjectifs n'ont pas reçu autant d'attention que les verbes et les noms bien qu'ils représentent une catégorie grammaticale très intéressante. En effet, les adjectifs ne présentent pas un ensemble homogène de propriétés.

Syntaxiquement, ils peuvent participer à différentes alternances, par exemple aux constructions attributives et prédicatives (par ex. *red* (rouge), *rectangular* (rectangulaire)) ou uniquement à l'une des ces constructions (par ex. *asleep* (endormi), *gastronomical* (gastronomique)). Ils peuvent avoir des compléments obligatoires (par ex. *reluctant* (réticent)), des compléments optionnels (par ex. *sad* (triste)) ou peuvent ne pas accepter du tout de compléments (par ex. *rectangular*). Certains adjectifs peuvent être mis au superlatif (par ex. *big* (grand), *quick* (rapide)), d'autres ne le peuvent pas (par ex. *finite* (fini), *carnivorous* (carnivore)). Sémantiquement, les adjectifs semblent se comporter différemment, certains adjectifs (par ex. *red*, *big*) sont extensionnels, c'est-à-dire qu'ils désignent des propriétés de premier ordre, d'autres (par ex. *alleged* (présumé), *former* (ancien)) sont intensionnels et représentent des propriétés de deuxième ordre. De nombreux essais théoriques ont été réalisés pour classifier les adjectifs. Certains se concentrent sur l'analyse de leur propriétés syntactiques (voir Huddleston, 1984; Quirk et al., 1985; Vendler, 1963, 1968), d'autres se concentrent sur l'analyse de leurs propriétés basées sur la théorie des modèles (voir Kamp, 1975; Keenan and Faltz, 1985; Keenan, 1987; Chierchia and Connell-Ginet, 1990; Kamp and Partee, 1995), et d'autres portent sur les distinctions ontologiques (voir Dixon, 1982, 1991; Frawley, 1992; Aarts, 1976). La classification des adjectifs qui est probablement la plus intéressante pour la reconnaissance d'implications est celle proposée par Kamp et Partee (Kamp, 1975; Kamp and Partee, 1995). Ils subdivisent les adjectifs anglais en quatre classes sémantiques tout en prenant en compte les modèles d'implication dont la source est sémantique et basée sur la théorie des modèles, et à laquelle participent les différents adjectifs. D'après cette classification, un adjectif tel que *red* est intersectif car la phrase [A N]$_{AP}$ implique la conjonction des propriétés exprimées par le nom N et par l'adjectif A. Un adjectif tel que *big* est subsectif car la phrase [A N]$_{AP}$ implique uniquement la propriété exprimée par N. Un adjectif tel que *fake* (faux) est privatif car la phrase [A N]$_{AP}$ implique la négation de la propriété exprimée par N. Et un adjectif tel que *alleged* est intégralement non-subsectif car rien ne peut être déduit de [A N]$_{AP}$. Cependant, aucune des classifications existantes ne peut entièrement décrire la complexité du comportement des adjectifs. La limite la plus fondamentale de ces classifications est qu'elles oublient de prendre en compte l'interaction entre les différents niveaux de la description linguistique, c'est-à-dire entre la syntaxe, la sémantique, la sémantique lexicale et la morphologie dérivationnelle, qui sont toutes des facteurs pertinents dans la détermination de possibles inférences. Comme les exemples (9) et (10) le montrent, bien que les deux adjectifs dry et rectangular peuvent être classifiés sémantiquement (sur la base de la théorie des modèles) comme étant intersectifs, ils présentent différents modèles d'inférence quant à l'antonymie, une propriété sémantique lexicale.

(9) a. *The dishcloth is dry → The dishcloth is not wet*
*(*Le torchon est sec → Le torchon n'est pas humide*)*

b. *The dishcloth is not dry → The dishcloth is wet*
*(*Le torchon n'est pas sec → Le torchon est humide*)*

(10) a. *This table is rectangular → This table is not round*
*(*Cette table est rectangulaire → Cette table n'est pas ronde*)*

b. *This table is not rectangular ↛ This table is round*
*(*Cette table n'est pas rectangulaire ↛ Cette table est ronde*)*

L'adjectif *dry* (sec) entre en relation antonymique binaire avec son antonyme *wet* (humide). C'est-à-dire que l'implication entre l'adjectif et la négation de son antonyme est valable dans les deux directions, c'est-à-dire $A \models \neg Anto(A) \wedge \neg A \models Anto(A)$. À l'inverse, l'adjectif rectangulaire ne possède pas un seul antonyme et entre dans de multiples relations oppositionnelles avec ses antonymes. Dans ce cas, la relation entre l'adjectif et la négation de l'un de ses antonymes est uniquement valable dans une direction, par ex. $A \models \neg Anto_1(A) \wedge \neg A \not\models Anto_1(A)$. Ceci laisse entendre que les quatre classes proposées par Kamp et Partee devraient être divisées en sous-classes et que des propriétés autres que celles purement basées sur la théorie des modèles devraient également être prises en considération. Par conséquent, dans cette thèse, j'individualise d'abord un ensemble d'aspects qui sont fondamentaux pour la description du comportement des adjectifs. Cet ensemble est obtenu en intégrant les critères de sélection utilisés dans les classifications d'adjectifs traditionnelles (Huddleston, 1984; Quirk et al., 1985; Vendler, 1963, 1968; Kamp, 1975; Keenan and Faltz, 1985; Keenan, 1987; Chierchia and Connell-Ginet, 1990; Kamp and Partee, 1995) avec des critères sémantiques lexicaux et morpho-dérivationnels. Puis, j'associe chaque propriété des adjectifs aux modèles d'inférence qu'elle génère et je classifie un échantillon de 500 adjectifs en utilisant ces modèles en tant que critères de sélection. Par conséquent, une classe sémantique est définie par l'ensemble de modèles d'inférence auxquels participent ses éléments. Il en résulte une classification sémantique à grain fin d'adjectifs anglais incluant 44 classes qui permettent de mieux prévoir leur comportement d'inférence. De plus, je développe un cadre sémantique compositionnel pour assigner à des adjectifs une représentation sémantique sur la base d'une approche ontologiquement variée (Hobbs, 1985) et qui permet ainsi l'inférence de premier ordre pour tous les types d'adjectifs, y compris les adjectifs extensionnels. Comme Hobbs (1985) et d'autres l'ont argumenté, la représentation sémantique de la langue naturelle ne doit pas nécessairement être d'ordre supérieur dans la mesure où la promiscuité ontologique peut résoudre le problème. C'est-à-dire qu'en réifiant tous les objets pouvant être décrits par un prédicat, il est possible d'obtenir un schéma de représentation sémantique de premier ordre pour la langue naturelle. Cette observation est cruciale pour les applications de traitement automatique des langues, et ce pour deux raisons. Premièrement, les logiques qui s'étendent au-delà du premier ordre sont extrêmement indécidables. Deuxièmement, et surtout, il n'existe pas de prouveurs automatiques usuels d'ordre supérieur pouvant être utilisés pour raisonner sur la signification de formules d'ordre supérieur. La représentation sémantique assignée aux adjectifs décrit également leurs propriétés sémantiques morpho-dérivationnelles et lexicales. Pour évaluer la classification proposée, je crée d'abord un corpus de test composé de paires de phrases illustrant les modèles d'inférence pertinents des adjectifs. Puis j'utilise le système de traitement automatique de la sémantique BB1 (Blackburn and Bos, 2005), premièrement pour associer automatiquement une représentation sémantique aux phrases du corpus de test, et deuxièmement pour tester si les membres de chaque paire de phrases sont ou non dans une relation d'implication. Les résultats montrent que la sémantique proposée prévoit correctement les modèles d'inférence observés pour les différentes sous-classes d'adjectifs identifiées dans cette thèse.

Conclusion

Dans cette thèse, j'ai étudié la manière dont les ressources lexicales basées sur l'organisation de la connaissance lexicales dans des classes qui partagent des aspects communs (syntactiques, sémantiques, etc.) permettent le traitement de la langue naturelle et en particulier la reconnaissance symbolique d'implications textuelles. Sur la base de la classification de verbes anglais par Levin (1993), j'ai tout d'abord présenté une approche robuste et à large couverture sur la reconnaissance de paraphrases. Puis, j'ai montré qu'en étendant le cadre proposé par Levin pour traiter les modèles d'inférence généraux, on obtient une classification d'adjectifs anglais qui, comparée à des approches antérieures, permet une caractérisation sémantique à grain plus fin de leurs propriétés déductives, et qui, par conséquent, constitue la base d'un traitement automatique de l'inférence basée sur les adjectifs. La principale contribution de ce travail repose dans une analyse détaillée des interactions entre la morphologie dérivationnelle, la sémantique lexicale et compositionnelle et de leur répercussion sur les modèles d'implication permis par des phrases contenant des adjectifs ou leur noms/verbes apparentés.

L'approche présentée ici offre la base d'un traitement automatique de l'inférence basée sur les adjectifs en fournissant une caractérisation à grain fin des différents types de modèles d'inférence permis par les adjectifs. Je crois que la méthodologie utilisée dans cette thèse pour classifier les adjectifs, c'est-à-dire la définition de classes sémantiques à l'aide du groupe de modèles d'inférence que permettent leur membres, est prometteuse pour la création d'ontologies générales basées sur la linguistique (Bateman, 1990) qui fournissent une interface syntaxe-sémantique pour les ontologies de connaissances nécessaires pour les outils du langage contrôlé. Un autre résultat notable est le corpus de test que j'ai développé en tant que ressource pour l'évaluation d'applications NLP qui traitent l'inférence (en particulier de l'inférence basée sur les adjectifs). Contrairement à d'autres corpus développés dans ce but, tels que les jeux de données du Paraphrase Research Corpus (C. et al., 2004; Dolan et al., 2004), du RTE1 (Dagan et al., 2006), du RTE2 (Bar-Haim et al., 2006) et du Text Retrieval Conference TREC (2007), ce corpus de test a une couverture clairement définie et contient des annotations pour les modèles d'inférence, ceci permettant un meilleur contrôle des données analysées. L'objectif de la construction de ce corpus de test est d'ouvrir la voie vers la création de ressources qui offrent une perspective approfondie dans les phénomènes linguistiques qui sont responsables de l'inférence.

Introduction

Understanding a text is one of the ultimate goals of computational linguistics. To achieve this goal, systems need to be developed which can construct a meaning representation for any given text and which, furthermore, can reason about the meaning of a text.

Recently, the PASCAL Recognising Textual Entailment Challenge RTE (Dagan et al., 2006) raised noticeable attention in the research community by focusing on one of the major inference tasks involved in textual reasoning, namely the entailment recognition task. This task consists in judging whether, given two texts T and H, called the hypothesis, the text T entails the hypothesis H.
For instance, in (1) the meaning conveyed by the text of the hypothesis (H), i.e. that *Joko Ono is the widow of John Lennon*, can be inferred from the information *her late husband, John Lennon* present in the text T, while in (2) the entailment between T and H does not hold as the information *Jupiter has four moons* does not follow from *four of Jupiter's tiniest moons*.

(1) T: *Joko Ono unveiled a bronze statue of her late husband, John Lennon, to complete the official renaming of England's Liverpool Airport as Liverpool John Lennon Airport.*

H: *Joko Ono is John Lennon's widow.*

(2) T: *They are made from the dust of four of Jupiter's tiniest moons.*

H: *Jupiter has four moons.*

The entailment recognition task has a wide range of applications, it can be used to determine whether a text fragment answers a question (in a question-answering application), whether a query is entailed by a relevant document (in information retrieval), whether a text fragment entails a specific information nugget (in information extraction), etc. As these applications for the most part focus on real text, the systems dealing with entailment recognition must be robust; that is, they must be able to handle unconstrained input. Most systems developed so far, therefore, are based on statistical methods (e.g. stochastic parsing and lexical distance or word overlap for semantic similarity) and few provide for a principled integration of lexical and compositional semantics.

On the other hand, one of the RTE-1 participant teams (Venderwende et al., 2005) has shown that roughly 50% of the entailment cases in the RTE-1 data set could be handled correctly by a system that would adequately cover semantic entailments that are either syntax-based (e.g. syntactic alternations) or lexical semantics-based (e.g. synonymy, antonymy). Given that the overall system accuracies hovered between 50 and 60 percent with a baseline of 50%, this suggests that a better integration of syntax, compositional and lexical semantics might improve entailment recognition accuracy. Further, the best results in the RTE-2 challenge were surprisingly achieved by two systems relying on the integration of deep and shallow techniques (Hickl et al., 2006) and on logical inference by means of handwritten inference rules and background knowledge (Tatu et al., 2006). Thus, proving for the first time that systems dealing with deep analysis perform better than statistic based ones.

However, the development of such systems and of high quality NLP systems in general, is made difficult by two factors which are closely related. First, there is a lack of appropriate wide coverage resources providing an integration of linguistic and background knowledge that is required to support reasoning and inference tasks. The construction of such lexical resources is indeed a very time-consuming task. Second, the lack of a deep understanding of the phenomena which originate entailment renders it often difficult to decide what kind of knowledge, i.e. whether linguistic or more general world knowledge, is actually required to deal with inference. Further, the notion of textual entailment has not yet been clearly defined. This is reflected in the difficulty to produce evaluation test beds for systems dealing with textual entailment. In the test data sets (e.g. Paraphrase Research Corpus[1], the RTE1[2], RTE2[3] data sets and the Text Retrieval Conference[4] data) appeared so far, the judgement of the entailment is not based on semantic equivalence but rather on a more "informal" notion of semi-equivalence, common sense, common human understanding and this often leads to disagreements and as some authors (Zaenan et al., 2005) have noticed even to errors.

I believe that a high quality computational treatment of textual entailment (i.e. of entailment problems that can be solved solely relying on the linguistic information which can be extracted directly from the text and not on world knowledge), whether statistic or symbolic, requires the precise definition of the notion of textual entailment as corresponding to logic entailment or semantic equivalence. Surely, this presupposes a deeper understanding of the linguistic phenomena responsible for (textual) inference and a theory of lexicon supporting the creation of lexical resources providing the knowledge necessary to deal with inference.

As some author noted, the approaches to lexicon theory which are more interesting for textual entailment and NLP in general are those based on the idea that *"the defini-*

[1]See (C. et al., 2004; Dolan et al., 2004)

[2]See (Dagan et al., 2006)

[3]See (Bar-Haim et al., 2006)

[4]See (TREC, 2007)

tion of a mapping between knowledge and its linguistic expression is facilitated if it is possible to classify any particular instances of facts, states of affairs, situations, etc., that occur in terms of set of general objects and relations of specified types that behave systematically with respect to their possible linguistic realisations" (Bateman, 1990). Indeed, lexical-semantic classifications of verbs have been proven useful in NLP (Dorr, 1997; Kipper et al., 2000; Klavans and Kan, 1998; Stede, 1998), as they allow generalisations and thus avoid redundancy in the lexicon. Further, these classifications support the automatic extraction of large-scale lexical resources from corpora (Brew and im Walde, 2002; Briscoe and Carroll, 1997; Merlo and Stevenson, 2001).

One of the most prominent work in this field is Levin's (1993) classification of English verbs. Based on the idea that syntactic and semantic behaviour are tight-knit, Levin defines verb classes modulo set of alternations. Thus all items in a verb class share a set of alternations, i.e. of paraphrastic constructions, while verbs in different classes differ in the paraphrastic constructions (alternations) they allow. So for instance, the verbs *meet* and *embrace* belong to different classes, as *meet* allows the *With Preposition Drop* alternation (3), while *embrace* does not (4).

(3) a. Anne met with Cathy

b. Anne met Cathy

*(4) a. *Brenda embraced with Molly*

b. Brenda embraced Molly

This idea is very promising for entailment recognition, and in particular for paraphrase recognition, as it permits to associate to two constructions that alternate the same semantic representation and further to automatically extract paraphrastic patterns for verbs, by only relying on the information available in the lexicon.

Thus, in this thesis, I investigate linguistically based approaches to recognition of textual entailment and paraphrases relying on fine grained semantics. In particular, I investigate how far lexical resources based on lexical semantic classifications of grammatical categories such as that of Levin can be extended in order to provide the knowledge required for dealing with inference. I focus on verb-based textual equivalence and adjectival entailment.

Thesis Contribution

In this thesis I make the following contributions:

1. I propose an approach to verb paraphrases which is linguistic-based. I demonstrate that it is robust and wide coverage by presenting its implementation in an incremental parsing framework.

2. I propose a methodology for a fine-grained semantic classification of English adjectives based on an extension of Levin's framework for verbs and based on test criteria which integrate formal semantics with lexical semantics, syntactic and morphoderivational information.

3. I present the classification of adjectives obtained by applying this method to 500 English adjectives.

4. I propose a flat event semantic representation for adjectives which builds on Hobbs' promiscuous ontologies. This formalism allows the representation of all types of adjectives, intensional and extensional, as first order predicates, thus making their computational treatment feasible. The formalism is also compositional and integrates the formal semantics of adjectives with their lexical semantics.

5. I then present a test suite for adjectival inference problems including about 3,000 test pairs.

Thesis Outline

This thesis is structured in two parts. The first part addresses the problem of verb-based textual equivalence in natural language and presents a linguistic-based approach to the textual equivalence recognition task. Chapter 2 introduces the linguistic problem of paraphrasing and reviews existing linguistic approaches to paraphrases.
Chapter 3 focuses on the lexical semantics of verbs. In this chapter, I present Beth Levin's classification of English verbs and I describe the type of information which can usefully be extracted from such a structured knowledge base to handle paraphrases. I then show how this knowledge can be expanded by also considering lexical knowledge about verbs coded in wide coverage resources such as WordNet.
Chapter 4 presents the proposed robust, linguistic-based approach to verb-based textual equivalence recognition which I developed based on Beth Levin's classification of English verbs and the XIP incremental parser.
Chapter 5 compares the proposed approach with related work on paraphrases.

The second part contains the main results of this thesis and presents a fine-grained semantic classification of English adjectives together with their semantic representation within a first order compositional framework. Chapter 6 gives a brief overview of previous attempts to classify English adjectives including syntactic and semantic classifications as well as attempts to organise adjectives in taxonomies. Chapter 7 presents the proposed fine grained semantic classification of adjectives based on an extension of Levin's framework as well as on adjectival features, ranging from syntax to lexical semantics and morpho-derivational properties, used as selectional criteria.
Chapter 8 specifies first order compositional semantics for adjectives by associating each adjectival class with a semantic schema capturing the compositional semantics of that class together with the set of axioms capturing their lexical semantic properties and presents the details of the framework within which the proposed approach to adjectival

inference has been evaluated.
Chapter 9 present the test suite for adjectival inference whose development is based on the results of the preceding chapters. Chapter 10 presents a comparison with existing computational lexical resources for adjectives. Chapter 11 concludes the thesis by discussing the results and by identifying directions for future research.

Part I

Verb Paraphrases

Paraphrases: Theoretical Background

A salient feature of natural language is that it allows paraphrases; that is, it allows different verbalisations of the same content.

Linguists have long noticed the pervasiveness of paraphrases in natural language and have attempted to characterise it. Thus, for instance, Chomsky's *transformations* capture the relation between one core meaning (a deep structure in Chomsky's terms) and several surface realisations (for instance, between the passive and the active forms of the same sentence) while Melcuk (1988) presents sixty paraphrastic rules designed to account for paraphrastic relations between sentences.
More recently, work in information extraction (IE) and question answering (QA) has triggered a renewed research interest in paraphrases as IE and QA systems typically need to be able to recognise various verbalisations of the content. Because of the large, open domain corpora these systems deal with, coverage and robustness are key issues and much of the work on paraphrases in that domain is based on automatic learning techniques.

In this chapter, I first introduce the notion of *paraphrase* and give a formal definition of it. I then present a typology of paraphrastic constructions which covers many of the cases dealt with throughout this thesis. Finally, I give a brief overview of linguistic theories which deal with paraphrases.

2.1 Paraphrases

Native language speakers are able to convey one and the same content by using very different verbalisations of it. Although these various verbalisations may have different pragmatic or communicative values (with respect for instance to topicalisation, presuppositions or focus/ground partitioning), they all share a core semantic content namely, the content approximated by a traditional montagovian compositional semantics.
For example, a native speaker of English can recognise that the sentences in (5)

(5) a. I loaded apples into the cart.

b. *I loaded the cart with apples.*

convey the same core meaning, namely that “the speaker has loaded apples into a cart”. This core meaning can be expressed by the formula

(6) `load(e)` $\wedge$ `agent(e,I)` $\wedge$ `patient(e,apples)` $\wedge$ `location(e,cart)`

In this thesis, I will take sentences to be paraphrases of each other whenever they share the same core meaning. Because their finer grained semantics and pragmatics usually differ, paraphrase can however differ in the entailments they license. Hence for instance, although (5a) and (5b) are paraphrases, they entail different things. Specifically, it can be inferred from sentence (5b) that after loading, the cart is full of apples whilst sentence (5a) only entails that after the loading event, the cart contains some apples.

Similarly, the sentences in (7) are also *paraphrases* in that they share the same core meaning even though they differ in emphasis.

(7) a. *Columbus discovered America.*

b. *America was discovered by Columbus.*

In sentence (7a), the emphasis is on *Columbus*, while in sentence (7b) it is on *America*. Now consider the sentences in (8), (9) and (10)

(8) a. *Columbus discovered America in 1492.*

b. *Columbus discovered America.*

(9) a. *The murderer slaughtered 10 people*

b. *The murderer killed 10 people*

(10) a. *The genome of the fugal pathogen that causes Sudden Oak Death has been sequenced by US scientists*

b. *The East Bay-based Joint Genome Institute said Thursday it has unraveled the genetic blueprint for the diseases that cause the sudden death of oak trees.*

c. *Scientists have figured out the complete genetic code of a virulent pathogen that has killed tens of thousands of California native oaks.*

In (8) and (9) the information expressed in one of the sentences of the pair is more detailed than that expressed in the other one:

$$\mathcal{C}ontent(S1) \subset \mathcal{C}ontent(S2) \qquad \textit{Entailment}$$

For example, the information about the point in time at which the event takes place (*in 1492*) is present in (8a) but absent in (8b). And in (9a), the verb *slaughter* specifies the way of the killing act expressed in (9b).
Finally, the sentences in (10) show content overlap:

$$\mathcal{C}ontent(S1) \cap \mathcal{C}ontent(S2) \cap \mathcal{C}ontent(S3) \neq \emptyset \qquad \textit{Overlap}$$

Cases such as (8), (9) and (10) of semantic entailment and overlap have been recently classified as paraphrases in QA and IE research. The Microsoft Research Paraphrases Corpus MSRP (Dolan et al., 2004) (the first corpus on paraphrases available on the web), for instance, considers sentences such as (10) as being paraphrastic despite the fact that they differ in content, as in MSRP the word *paraphrase* is used to characterise *semantic near equivalence*. And Glickman and Dagan (2003) classify cases such as (8) and (9), in which the meaning of one member of the pair (specifically a verb) entails the meaning of the other (but not vice versa), as paraphrases.

Since the word *paraphrase* is used differently in the computational linguistic literature, I now give a formal definition of how this word should be interpreted throughout this thesis before presenting a classification of paraphrastic means.
In the following, I regard differences in communicative value (e.g. emphasis) as being irrelevant to paraphrasing, while differences in content will be considered relevant. Thus, in the context of this thesis, the word *paraphrase* means *textual equivalence*[5] Two sentences are *textually equivalent* iff the entailment relation between them is bidirectional. Thus, the task of recognising textual paraphrases corresponds to the task of deciding about the *textual equivalence* of two sentences.

> Given two sentences $\mathcal{S}_1$ and $\mathcal{S}_2$, it holds that $\mathcal{S}_1$ is a *paraphrase* of, or is *textually equivalent* to $\mathcal{S}_2$ (written $\mathcal{S}_1 \leftrightarrow \mathcal{S}_2$), iff $\mathcal{S}_1$ entails $\mathcal{S}_2$ (written $\mathcal{S}_1 \rightarrow \mathcal{S}_2$) and $\mathcal{S}_2$ entails $\mathcal{S}_1$:
>
> $$\mathcal{S}_1 \leftrightarrow \mathcal{S}_2 \text{ iff } \mathcal{S}_1 \rightarrow \mathcal{S}_2 \wedge \mathcal{S}_2 \rightarrow \mathcal{S}_1.$$

The notion of textual entailment corresponds here to the notion of logic entailment between the representations of the two sentences. Thus, the definition above can be rewritten as follows

(11) $\mathcal{S}_1 \leftrightarrow \mathcal{S}_2$ *iff* $\Phi(\mathcal{S}_1) \models \Phi(\mathcal{S}_2) \wedge \Phi(\mathcal{S}_2) \models \Phi(\mathcal{S}_1)$.

where $\Phi(\mathcal{S}_i)$ corresponds to the logic representation of the sentence S_i.
This definition prevents sentence pairs such as (8), (9) and (10). Sentence pairs such as (5) will be deemed paraphrases or not depending on how fine grained their representation is. If they are both assigned the semantic representation sketched above, they will be deemed paraphrases. If however, they are assigned a more fine grained semantic representation reflecting their aspectual difference, they will not.

[5] Although in this thesis, I am mainly concerned with sentences, I will follow the general trend induced by the RTE (Recognising Textual Entailment) challenge and speak of *textual* (rather than sentential) entailment/equivalence.

2.2 Textual Equivalence Based on Domain-Independent Lexical and Structural Knowledge

Natural language makes available numerous means to produce paraphrastic structures. Reformulations of the same meaning can have different sources depending on the kind of knowledge which is the source of textual equivalence. I will distinguish between *lexico-structural paraphrases*, which are those for which the source of textual equivalence is lexical knowledge or knowledge about the syntactic relations involved in the pair of paraphrases, and *extra-linguistic paraphrases*, which are those in which the source of the equivalence is world knowledge or the context of utterance.
The sentences in (5), for instance are lexico-structural paraphrases as the nature of the relations between the two sentences can be explained in terms of different syntactic realisations. But consider example (12) below where a referential expression is substituted with its referent.

(12) a. *I wrote this letter two months ago.*

b. *I wrote this letter in August.*

In this case, lexical semantics are not enough to infer that the two sentences are paraphrases. Knowledge about the context and the semantics of the referent expression is necessary.
Similarly, the sentences in (13) are clearly paraphrases, but again the source of semantic equivalence is neither lexical nor structural. In this case world knowledge is used to establish semantic equivalence, namely the knowledge that newborn rabbits are to young to have teeth.

(13) a. *I found a newborn rabbit in the garden.*

b. *I found a newborn rabbit with no teeth in the garden.*

In this thesis, I focus on lexico-structural textual equivalence. Specifically, on paraphrases where semantic equivalence relies either on lexical or on structural equivalence between natural language expressions. I now review each of these two types of paraphrases in more detail.

2.2.1 Lexical Paraphrases

Lexical paraphrases are those cases of textual equivalence which can be solved by relying solely on domain-independent lexical knowledge found in the lexicon and which do not involve any change in the syntactic structure of the sentence. *Synonymy* and *antonymy* are typical exponents of this type of paraphrastic constructions.

Synonymy

In the case of synonymy, semantic equivalence is obtained by means of substitution of two lexical units sharing the same meaning. Depending on the type of the lexical units involved, synonymy can be further divided into single- and multi-word synonymy.

Single-word synonymy This type of synonymy involves cases such as (14) where the lexical unit which should be substituted and its synonym are single words.

(14) a. John is playing the violin.

b. John is playing the fiddle.

Multi-word synonymy In this case, a semantic equivalence holds between a lexical unit and a phrase defining the meaning of the lexical unit.
In some cases, derivational morphology can help to derive the equivalence relations between word and phrase, as in (15) and (16).

(15) a. The driver of the BMW is bald.

b. The person who drives the BMW is bald.

(16) a. His skin reddened.

b. His skin became red.

But often, multi-word synonymy does not involve derivational morphology and the meaning of a lexical unit is defined through of a complex phrase. For instance in (17), the phrase *increase the number of* expresses the meaning of *proliferate* and in (19) the meaning of *dissipated* is expressed in terms of a conjunction *spread and disappeared*.

(17) a. The senators proliferated government subsidies.

b. The senators increased the number of government subsidies.

(18) a. The bones fossilised.

b. The bones got preserved in stone.

(19) a. The smoke dissipated.

b. The smoke spread and disappeared.

(20) a. Who was president in 1881?

b. Who was head of state in 1881?

In order to handle such kind of paraphrases, definitional information (i.e. world knowledge information) needs to be included in the lexicon.

Antonymy

In the case of antonymy, semantic equivalence is obtained by means of substitution of two lexical units sharing the same meaning modulo negation. Again, as for synonymy, antonymy can be realised between single words (single-word antonymy) or between a word and a phrase (multi-word antonymy).

Single-word antonymy This type of antonymy involves cases such as (22) where a single word is substituted with the negation of its antonym[6] (also a single word).

(22) a. *I remember that picture.*

b. *I haven't forgotten that picture.*

Multi-word antonymy This type of antonymy involves cases such as (23) where a lexical unit is substituted with the negation of the definition of the lexical unit corresponding to its antonym.

(23) a. *John is a man.*

b. *John is not a person of female gender.*

Again, as for multi-word synonymy, capturing such types of antonymy can require definitional information.

2.2.2 Structural Paraphrases

Structural paraphrases are those cases of textual equivalence which can be solved by considering the syntactic properties of the lexical units involved in the text. Typical examples of such cases of paraphrastic constructions are *syntactic variations, alternations* and *converse constructions*.

[6]The context in which the substitution happens has a great importance as negation may interact with other operators. As example (21) shows, the semantic equivalence between a word and the negation of its antonym no longer holds if the substitution happens in a question context.

(21)

Did you remember his name?

Did you not forget his name?

In this case, negation takes scope over the question operator so that the two questions can be paraphrased as *is it true that you remembered his name?* and *is it not true that you forgot his name?*. A proper treatment of such cases would involve a treatment of scope which I will not go into here.

Grammatical Variations

Anderson (1971) is perhaps the first author who, within the transformational grammar theory, makes a distinction between *grammatical variations* which are common to all items of a grammatical category and can be captured by a general rule (in his case a transformational rule), and syntactic variations which cannot, namely *alternations*.

The first kind of variation (grammatical variation) does not depend on the semantic class the lexical item belongs to. To this type of variations belong among others it-cleft constructions, topicalisation, question and relative clause formation.

(24) a. *John left the town.* *it-cleft*

b. *It was John who left the town.*

(25) a. *The Bahamas, you said, were warm in January.* Topicalisation

b. *You said the Bahamas were warm in January.*

(26) a. *The officials are honest.* Questions

b. *Are the officials honest?*

(27) a. *John saw a man running.* Relative clause

b. *John saw a man who was running.*

Alternations

Alternations relate syntactic realisations of the same (or of a similar) lexical meaning. However, they are much less regular and are constrained by the semantic class to which the lexical unit belongs. So, for instance, not all verbs present the *dative* alternation. Alternations can be further subdivided into *intra*- and *extracategorial alternations* depending on whether the syntactic variations under consideration involve a change in the syntactic category of the lexical item which originates it or not.

Intracategorial alternations In this type of variation, semantic equivalence is obtained by rearranging the linking between grammatical functions and thematic roles of the arguments of a predicative lexical unit.

(28) a. *John gives Ann the book.* Dative Alternation
John gives the book to Ann.

b. *This key opens the safe.* Instrument Subject Alternation
The safe opens with this key.

c. *The laboratory merges with the firm.* Simple Reciprocal Alternation Intr.
The laboratory and the firm merge.

The sentence pairs in (28) are examples of verb-based alternations. Figure 2.1 illustrates the dative alternation involved in (28a), i.e. the two mappings between roles and grammatical functions possible for verbs which allow this alternation. So for instance, the agent role (or subject) of the verb *give* is realised in both sentences as NP in nominative (*John*), the theme role (or direct object) of the verb *give* is realised as NP in accusative (*book*), while the benefactive role (or indirect-object) of the verb can be syntactically realised in two different ways, as an NP in dative (*Ann*) or as a PP (*to Ann*).

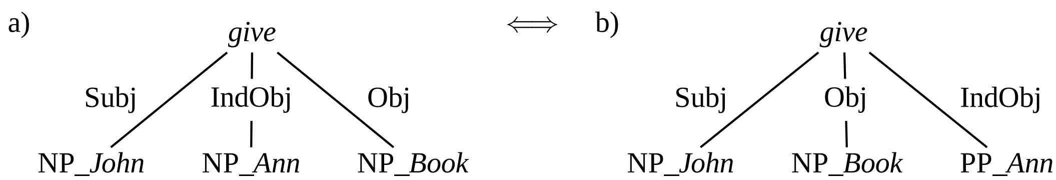


Figure 2.1: Dative Alternation

Intracategorial alternations can also involve non verbal predicates, for instance adjectives as in (29).

(29) a. *It is certain that the mayor will chastise the invaders for undermining his authority.*

b. *The mayor is certain to chastise the invaders for undermining his authority.*

Extra-categorial alternations In such cases, semantic equivalence results from modifying the syntactic category of a predicative lexical unit while preserving, or not, the mapping between grammatical functions and semantic roles. In (30) and (31), for instance, the semantic equivalence is obtained by substituting a verb with a derivationally related noun expressing the same meaning, while in (32) a noun is substituted with a semantically related adjective.

(30) a. *Smith is the inventor of the process.*

b. *Smith invented the process.*

(31) a. *The invention of the process is due to Smith.*

b. *Smith invented the process.*

(32) a. *A is 2 km away from B.*

b. *The distance between A and B is 2 km.*

Converses

In (33) and (34), semantic equivalence is realised by using *converse* constructions.

(33) a. *Ann gives books to John.*

b. *John receives books from Ann.*

(34) a. *Ann is John's mother.*

b. *John is Ann's son.*

In such constructions, a predicative lexical unit is substituted with the *converse* predicative lexical unit which has same meaning, i.e. describes the same situation, and same arity but has an opposite or 'converse' linking of grammatical functions and thematic roles.

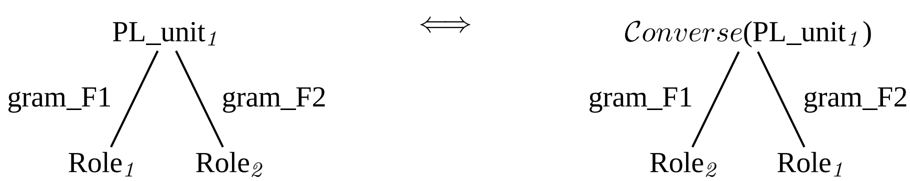


Figure 2.2: Converse Constructions

Converses of lexical items, like synonyms, should be registered in the lexicon in order to be recognised and used for paraphrasing.

2.2.3 Combination of Different Paraphrastic Means

Natural language allows for more complex cases of textual equivalence which originate from the combination of previously described paraphrastic means.

(35) a. *Caesar was killed by Brutus.*

b. *Brutus was the murderer of Caesar.*

c. *The death of Caesar was caused by Brutus.*

Consider the sentences in (35). As illustrated in Figure 2.3, the detailed list of paraphrastic equations which applied to sentence (35a) give as a result the sentences (35b) and (35c), the second and third sentence can be derived from the first by applying in sequence some of the previously described paraphrastic means.

Caesar was killed by Brutus. ≡ *Brutus was the murderer of Caesar.*	
Caesar was killed by Brutus.	
≡ *Brutus killed Caesar.*	*active/passive variation*
≡ *Brutus murdered Caesar.*	*synonymy*
≡ *Brutus was the murderer of Caesar.*	*extra-categorial alter.*

Caesar was killed by Brutus. ≡ *The death of Caesar was caused by Brutus.*	
Caesar was killed by Brutus.	
≡ *Brutus killed Caesar.*	*active/passive variation*
≡ *Brutus caused the death of Caesar.*	*definition*
≡ *The death of Caesar was caused by Brutus.*	*active/passive variation*

Figure 2.3: Examples of combinations of paraphrastic means

2.3 Linguistic Work on Paraphrasing

I will now present some of the most prominent linguistic approaches to paraphrasing.

2.3.1 Transformational Grammar

One of the first attempts to give a principled account of paraphrasing in a linguistic theory is probably Chomsky's system of transformational grammar (Chomsky, 1957), developed in the 1960's by building on the work of Harris (1951). In transformational grammar theory, a sentence has two levels of representation: the *deep structure*, represented as a phrase structure tree, and the *surface structure*. A system of formal rules describes the relation between one deep structure and its possible different verbalisations as surface structure.

The meaning of a sentence is given in the deep structure, thus if two sentences have the same deep structure, then they must be paraphrases. For instance, consider the paraphrastic sentences (36).

(36) *a. The policeman killed a demonstrator.*

b. A demonstrator was killed by the policeman.

c. There was a demonstrator killed by the policeman.

These sentences can be assigned the (simplified) syntactic structures shown in Figure 2.4. Transformational grammar theory can predict that these three sentences are paraphrases, as their phrase structure tree representations can be transformed one in the other by consecutively applying the transformation rules for active-passive and there-insertion illustrated in Figure 2.5.

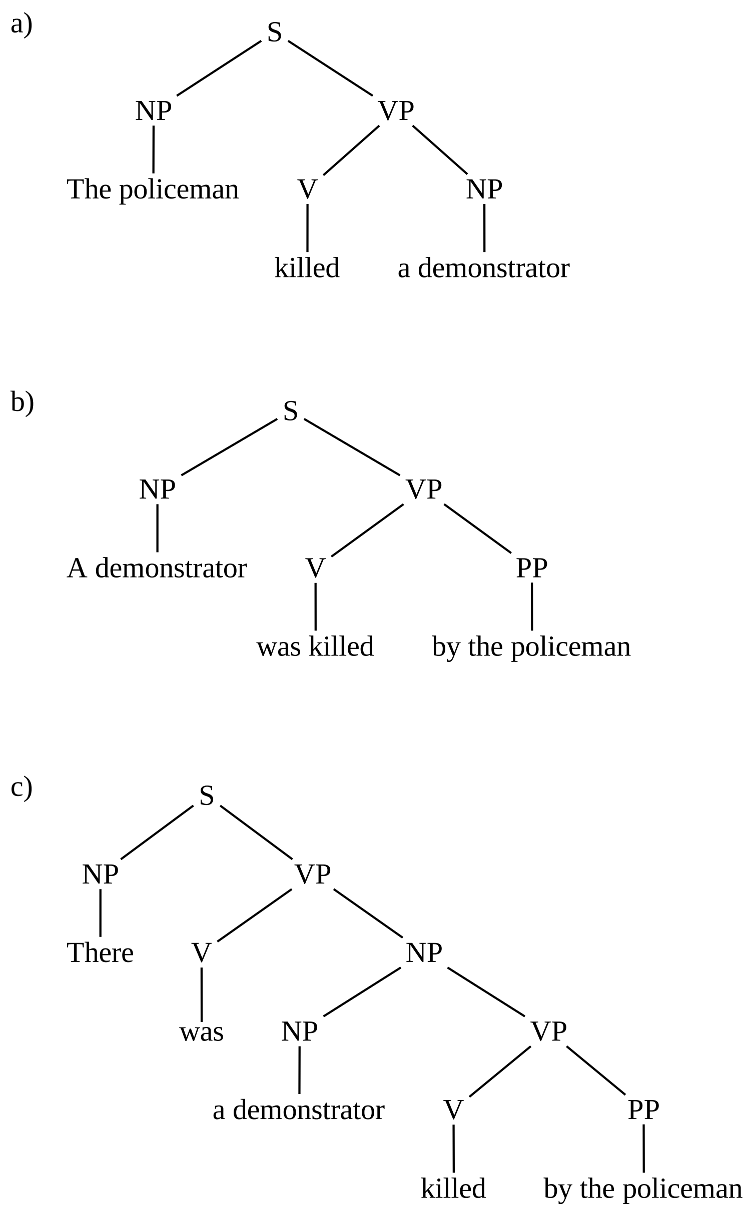


Figure 2.4: Active-Passive transformation followed by There-insertion

The *Active-Passive* transformation rule, for instance, applies to any structure in which the terminal string, i.e. the sentence, is analysable as a noun phrase NP_1 immediately followed by a verb V which is immediately followed by a noun phrase NP_2 (it does not matter what constituents (X) precede or (Y) follow the two noun phrases). In the new structure resulting after application of the rule, the verb is preceded by an auxiliary verb used to build its passive form, the two NPs (NP_1 and NP_2) are inverted so that the originary subject-NP (NP_1) is preceded by the preposition *by* and corresponds to the indirect object of the verb.

Active-Passive Transformation Rule

$$\begin{array}{l} X - NP_1 - V - NP_2 - Y \\ 1 \quad 2 \quad\; 3 \quad 4 \quad\; 5 \quad ====> \quad 1, 4, Aux, 3, by, 2, 5 \end{array}$$

There-Insertion Rule

$$\begin{array}{ccccl} \text{NP[-def]} - & \left\{ \begin{array}{r} be \\ exist \\ arise \end{array} \right\} & - X & & \\ 1 & 2 & 3 & ====> & \text{There, 2, 1, 3} \end{array}$$

Figure 2.5: Transformation Rules

The *There-Insertion* rule inserts the string *there* in the phrase structure tree of a sentence in which the subject-NP is a NP with undefined article ([-def]) immediately followed by a verb belonging to the set of verbs of existence (e.g. *be*, *exist*, *arise*, etc.).

The Transformational Grammar Theory approach, although interesting for the treatment of paraphrases, deals solely with grammatical variations and does not present any analysis of the more complex alternations. Moreover, it is not declarative; no clear distinction is made between grammar and lexicon. And, more importantly for the scope of this thesis, no computational implementation has been developed so that it is not possible to give an evaluation of the ability of this approach to handle paraphrasing.

2.3.2 Meaning-Text Theory

Probably the most important and comprehensive treatment of paraphrases was developed within the *Meaning-Text Theory* (Melcuk and Zolkovskij, 1970, 1984), a theory explicitly designed for translation and paraphrasing.

Melcuk (1988) identifies the description of a natural language with the construction of a system of formal rules which, given a meaning, enables the generation of a text conveying this meaning and of all its paraphrases. The Meaning Text Theory is a *stratificational* model of language which has three components:

- a structured representation of the text

- a set of translation rules
- a lexicon

The structured representation of the text consists of seven levels (see Figure 2.6) capturing semantic, syntactic, morphological and phonological aspects of the text.

SemR Semantic Representation
$\Updownarrow$
DSyntR Deep-syntactic Representation
$\Updownarrow$
SSyntR Surface-syntactic Representation
$\Updownarrow$
DMorphR Deep-morphological Representation
$\Updownarrow$
SMorphR Surface-morphological Representation
$\Updownarrow$
DPhonR Deep-phonological Representation
$\Updownarrow$
SPhonR Surface-phonological Representation

Figure 2.6: Levels of text representation in Meaning-Text Theory

- The semantic representation SemR is a decompositional representation of the meaning of the text (or of the lexical item) based on a number of semantic primitives. This representation corresponds to a semantic network, a connected directed graph in which nodes represent semantic primitives and arcs semantic dependencies labelled with numbers. Figure 2.7, for instance, shows the MTT semantic representation level SemR for the preposition *before*. The lexical meaning of the preposition is given in abbreviated form in a), which shows that the preposition *before* has two semantic actants or arguments 'X' and 'Y'. The decompositional representation correspondent to a) is given in b), where the meaning of the preposition is represented in terms of the semantic primitives `time_of_event` and `greater_than`. The representation in b) can be read as follows: the preposition *before* has two semantic actants or arguments,'X' and 'Y', which represent the two events which happen respectively at points in time t1 and t2, where t1 precedes t2.

- The deep syntactic representation DSyntR of a text (see Figure 2.8 a) corresponds to a sort of flat semantic representation, actually a dependency tree whose nodes

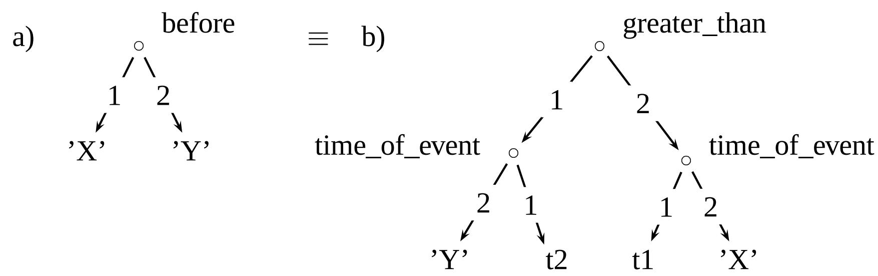


Figure 2.7: MTT Semantic Representation Level SemR of the preposition '*before*'

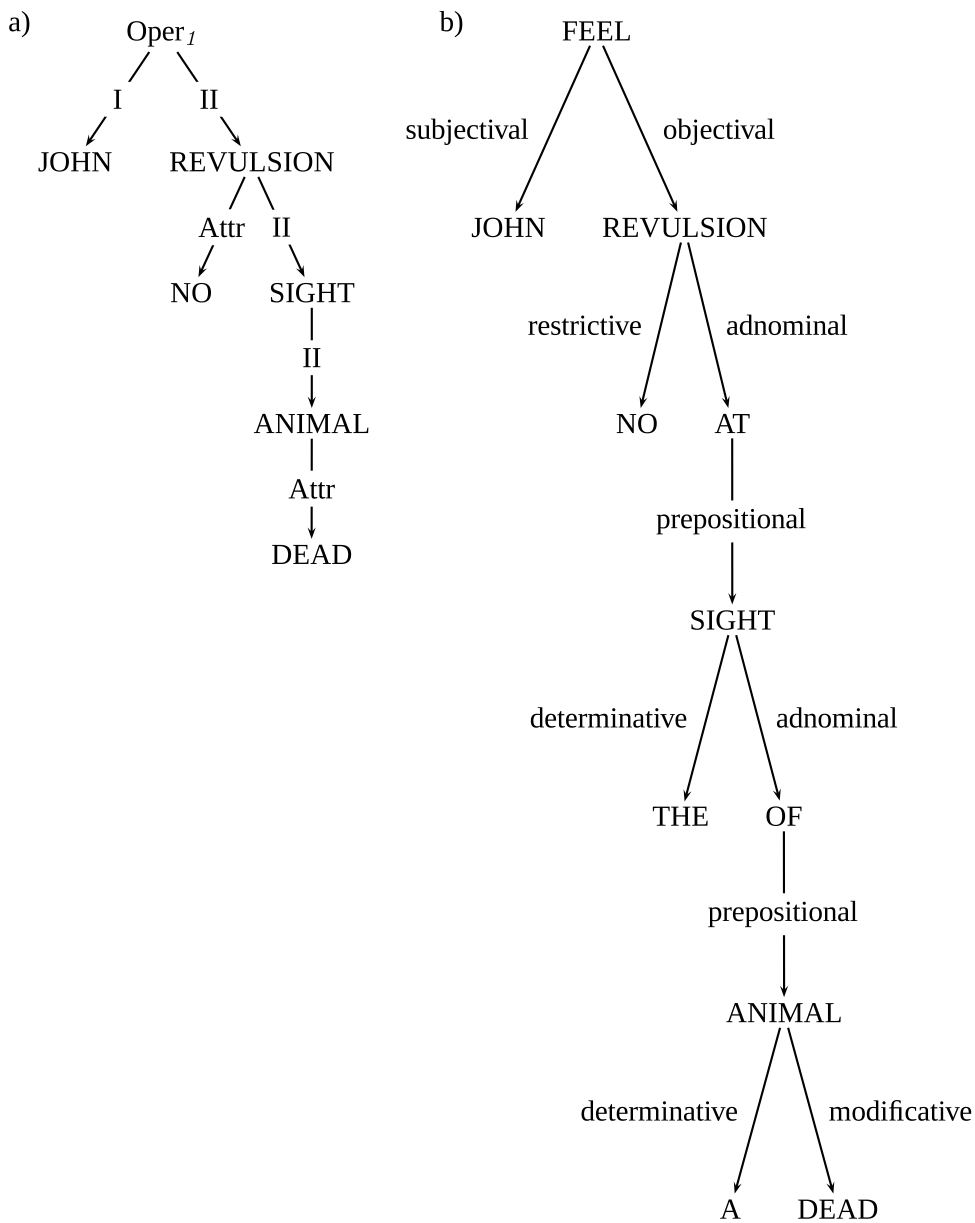
are not linearly ordered. The nodes represent semantically full lexemes (semantic empty lexemes such as conjunctions, auxiliary verbs, prepositions etc. are represented in the next level SSyntR).

- The surface syntactic representation SSyntR (see Figure 2.8b) is also a dependency tree representing the syntactic constituency of the sentence.
- The deep morphological representation DMorphR (see Figure 2.8c) corresponds to an ordered list of lexemes labelled with their corresponding morphological and agreement information.
- The surface morphological representation SMorphR (see Figure 2.8d) of a word form corresponds to the set of morphemes which constitute it.
- The deep phonological representation DPhonR introduces information about phonological properties of the words.
- Finally, the surface phonological representation SPhonR gives the real prosodic structure of the sentence.

The set of translation rules describes the transformation needed to make a transition from a level of representation to the following one. A *transition* or *correspondence* rule has the form

$$X \Longleftrightarrow Y \mid C,$$

where X and Y are representations of two adjacent levels and C is a constraint on the transition. Such a rule is read as follows: *if condition C is verified, X representation can be translated into Y representation and vice versa.* Figure 2.9 shows an example of a translation rule from a deep-syntactic representation to a surface-syntactic representation.
This rules applies to a segment of the text whose deep syntactic representation corresponds to a situation in which Y is the first syntactic actant of the verb X. If X is a finite verb, then such a deep-syntactic representation can be translated in the surface form Y + X where Y is the subject of the verb X.



c) [John$_{sg}$ Feel$_{ind,pres,3,sg}$ No Revulsion$_{sg}$ At The Sight$_{sg}$ Of Dead Animal$_{pl}$]

d) [John$_{sg}$ Feel+s$_{ind,pres,3,sg}$ No Revulsion$_{sg}$ At The Sight$_{sg}$ Of Dead Animal+s$_{pl}$]

Figure 2.8: MTT Representation of the sentence '*John feels no revulsion at the sight of dead animals*.'

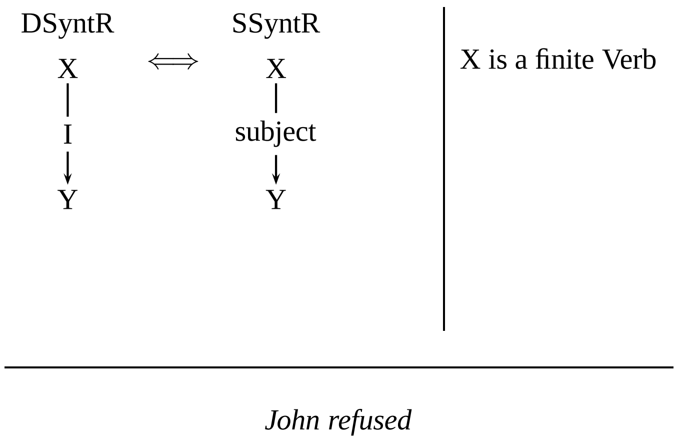
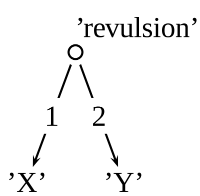


Figure 2.9: DSyntR-Level to SSyntR-Level Transition Rule

The lexicon, the Explanatory-Combinatorial Dictionary (ECD), provides language-specific information. It contains very exhaustive syntactic and semantic information about lexical units. This information is referenced by the various transition rules. An ECD entry has the following component:

- *semantic component*, a decompositional representation of the meaning of the lexical item in propositional form describing the corresponding semantic network. Figure 2.10 b gives as an example the semantic representation of one sense of the noun *revulsion* (which is abbreviated as in Figure 2.10 a). The symbol | separates the presuppositional part of the definition (on the left) from the assertional (on the right side).

a)



b) `Xs revulsion for Y :` X perceiving Y,| Xs (strong) negative emotion about Y which is similar to what people normally experience when they are in contact with something that makes them sick and such that it causes X to want to avoid any contact with Y.

Figure 2.10: Example of EDC semantic component

- *syntactic component*, a specification of all surface syntactic means with which it is possible to express the semantic arguments of the lexical unit. The syntactic component for the noun *revulsion* shown in Figure 2.11 gives all the possible syntactic realisations for the semantic arguments of the noun listed in the semantic component. The X semantic argument of the noun, for example, is always realised

X	Y
1. N	1. *against* N 2. *at* N 3. *for* N 4. *toward* N

Figure 2.11: Example of EDC syntactic component

as a noun, while the Y component may be realised in different ways. Thus, the information represented in Figure 2.11 describes, among others, the following different syntactic patterns possible for the noun *revulsion*:

(37) *John's revulsion against racism/at Mary's greed.*
John's revulsion at such behaviour/at the sight of seafood.
John's revulsion for work/ for all those killings.
John's revulsion toward all those scoundrels/toward the government.

- *lexical component*, a specification of the lexical functions of the linguistic unit. Through lexical functions, the MTT represents lexical semantic, morphological and syntactic properties of lexical items, such as, for example, synonymy (Syn), antonymy (Anti), converses (Conv_{ij}), nominalisation (Si), morphologically related adjectives (A_i), collocations (Oper_i, Func_i and Labor_{ij}), etc. Figure 2.12 shows the ECD lexical component for one of the senses of the noun *revulsion*.

Paraphrasing and Translation in MTT

In order to capture lexical and structural paraphrases, Melcuk (1988) defines *equivalence* rules which apply at the deep-syntactic level.

> If a semantic representation SemR_A corresponds to two different deep-syntactic representations DSyntR_{B_1} by rule R1 and DSyntR_{B_2} by rule R2, then there exists an equivalence rule or a paraphrasing rule R_{par} between DSyntR_{B_1} and DSyntR_{B_2} (see Figure 2.13).

Melcuk (1988) individuates about sixty equivalence rules for paraphrasing which are defined for French but he argues that these are also applicable to any other language. Some of these equivalence rules are illustrated in Figure 2.14.
A paraphrasing rule is composed of two parts:

Syn	: distaste
Syn_i	: repugnance
Anti_i	: attraction
$\text{Conv}_{21}\text{Anti}_i$	: appeal_N
A_1	: revulsed
Magn + A_1	: be filled [with ∼ (about N=Y)]
Adv_1	: in[∼]
Magn + Adv_1	: well up in [∼]
Propt	: from [∼]
Able_2	: revulsive
Magn + Able_2	: of utmost [∼] | G[7]= scene, sight
Qual_1	: squeamish; overly sensitive
Magn	: violent $<$ extreme $<$ utmost
AntiMagn	: slight
Oper_1	: experience, feel [∼ for/toward N=Y]
$\text{Conv}_{21}\text{Caus}_2\text{Oper}_1$	: be driven [to ∼] | passive only
Magn + Labor_{21}	: fill [N=X with ∼]
Caus_2	: revolt [N=X]
Adv_1Manif	: with [∼]

Figure 2.12: Lexical semantics of the noun *revulsion*

- a lexical rule, stating an equivalence relation between a lexical unit C_0 and a composition of some of the lexical functions defined for such a unit in the lexicon. For instance, the lexical unit *daughter*(C_0) is related to the lexical unit *parent* through the lexical function $Conv_{21}$, i.e. *parent* is equivalent to $Conv_{21}$ applied to *parent*.

- a syntactic rule, describing the syntactic transformation at the surface level after application of the lexical rule. For instance, after application of the transformation *daughter*(C_0) $\rightarrow$ *parent*($Conv_{21}(C_0)$), the surface syntactic position of the two syntactic actants (actant I = the parent, actant II= the daughter) is exchanged.

These rules predict paraphrastic patterns such as those in (38).

(38) a. *John plays the violin*$[C_0]$. *Rule 1*
John plays the fiddle$[Syn(C_0)]$.

b. *John*$[I]$ *is the parent*$[C_0]$ *of Mary*$[II]$. *Rule 2*
Mary$[II]$ *is the daughter*$[Conv_{21}(C_0)]$ *of John*$[I]$.

c. *John*$[I]$ *gives*$[C_0]$ *a book*$[II]$ *to Mary*$[III]$. *Rule 4*
Mary$[III]$ *receives*$[Conv_{321}(C_0)]$ *a book*$[II]$ *from John*$[I]$.

[7] G stands for the Syntactic Governor of the value of the Linguistic Function in question.

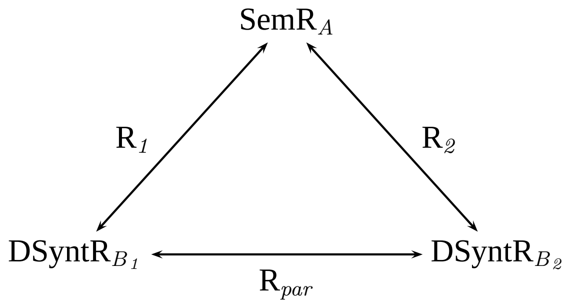


Figure 2.13: Definition of Equivalence Rule

d. *John is fast*$[C_0]$. *Rule 17*
John is not slow$[Anti(C_0)]$.

e. *Her sorrow*$[S_0(C_0)]$ *is deep*$[Pred(A)]$. *Rule 36*
She suffers$[C_{0(v)}]$ *deeply*$[A]$.

Rule 1, for example, describes lexical synonymy: a lexical item C_0 (violin) is substituted with its synonym $Syn(C_0)$ (fiddle). Rule 4 describes a converse construction: a verbal lexical item $C_{0(V)}$(give) is substituted with its converse $Conv_{21}(C_0)$ (receive) while the semantic arguments of the verb are inverted. And rule 36 describes a paraphrastic pattern in which a noun plus attribute construction (deep sorrow) is substituted with a verb (suffer) morphologically or semantically related to the noun plus the adverb (deeply) which is morphologically or semantically related to the adjective.

The Meaning-Text Theory permits a comprehensive treatment of paraphrases but has important shortcomings. First, the construction of an ECD lexicon for a language is very time consuming. ECDs for Russian (Melcuk and Zolkovskij, 1984) and French have been developed but not for English. Second, there is no wide coverage parser for English using all levels of representation of the MTT theory.

2.4 Summary

In this chapter, I discussed the importance of paraphrasing in natural language and gave a typology of paraphrases that can be captured by solely considering domain-independent lexical knowledge.
I then gave an overview of the most important approaches to paraphrasing in linguistic theory, namely the transformational grammar approach and the Meaning-Text Theory approach. Such linguistic theoretical approaches clearly show the importance of lexical information to handle paraphrases but also present some important drawbacks. Transformational grammar theories have limited coverage (they only handle syntactic variations) and no computational implementation. Melcuk MTT presents a detailed and exhaustive treatment of paraphrases which relies on a very rich lexicon, but it is very difficult to

Lexical Rules	**Syntactic Rules**
1. $C_0 \Longleftrightarrow Syn(C_0)$	1. $X \xrightarrow{I} A,\ X \xrightarrow{II} B \Leftrightarrow Y \xrightarrow{I} A,\ Y \xrightarrow{II} B$
2. $C_{0(V)} \Longleftrightarrow Conv_{21}(C_0)$	19. $X \xrightarrow{I} A,\ X \xrightarrow{II} B \Leftrightarrow Y \xrightarrow{II} A,\ Y \xrightarrow{I} B$
4. $C_{0(V)} \Longleftrightarrow Conv_{321}(C_0)$	20. $X \xrightarrow{I} A,\ X \xrightarrow{III} C \Leftrightarrow Y \xrightarrow{III} A,\ Y \xrightarrow{I} C$
17. $C_{0(V)} \Longleftrightarrow Anti(C_0)$	1. $X \xrightarrow{I} A,\ X \xrightarrow{II} B \Leftrightarrow Y \xrightarrow{I} A,\ Y \xrightarrow{II} B$
18. $C_{0(V)} \Longleftrightarrow S_0(C_0) \overset{II}{\leftarrow} Oper_1(S_0(C_0))$	11. $X \xrightarrow{II} B,\ X \xrightarrow{III} C \Leftrightarrow Y \xrightarrow{II} B \xrightarrow{II} c$
36. $C_{0(V)} \overset{ATTR}{\rightarrow} A \Longleftrightarrow \left\{ \begin{array}{c} S_{mod}(C_0) \\ S_0(C_0) \end{array} \right\} \overset{I}{\leftarrow} Pred(A)$	25. $Z \xrightarrow{I} Y,\ Z \xrightarrow{II} W \Leftrightarrow X \xrightarrow{ATTR} W$
40. $Oper_1(C_0) \Longleftrightarrow Func_1(C_0)$	19. $X \xrightarrow{I} A,\ X \xrightarrow{II} B \Leftrightarrow Y \xrightarrow{II} A,\ Y \xrightarrow{I} B$
48. $CausX(C_{0(N)}) \Longleftrightarrow IncepX(C_{0(N)})$	19. $X \xrightarrow{I} A,\ X \xrightarrow{II} B \Leftrightarrow Y \xrightarrow{II} A,\ Y \xrightarrow{I} B$
49. $CausX(C_{0(V)}) \Longleftrightarrow X(C_{0(V)})$	19. $X \xrightarrow{I} A,\ X \xrightarrow{II} B \Leftrightarrow Y \xrightarrow{II} A,\ Y \xrightarrow{I} B$

Figure 2.14: Examples of MTT Paraphrasing Rules

realise in practise and lacks computational implementation.

In the next chapter, I follow up these ideas and present a linguistic theory along with computational wide coverage resources which support a computational treatment of some paraphrases, namely verb-alternation paraphrases.

Linguistic Resources for Verb Paraphrasing

As shown in the previous chapter, linguistic-based approaches to paraphrasing commonly focus on finding a system of rules describing paraphrastic patterns. Paraphrasing rules directly operate on linguistic information which is stored in a lexicon containing detailed information about the syntax and the semantics of lexical items. However, the creation of such a lexicon in practise is prevented by the problem that it is very time consuming.

Nevertheless, lexical resources based on the organisation of lexical knowledge in classes which share a common linguistic behaviour, such as the classification of English verbs proposed by Levin (1993), result useful as they allow to reduce the redundancy of the lexicon and provide a syntax semantic interface for NLP systems. For paraphrasing, a classification such as Levin's is particularly interesting as, being based on the definition of alternations, i.e. structural paraphrases, it also allows the automatic extraction of a set of rules describing paraphrastic patterns for verbs. Such paraphrastic patterns can be further extended by considering the linguistic information (e.g. synonymy/antonymy, hyponymy etc.) encoded in resources such as WordNet (Fellbaum, 1998).

In this chapter, I first present Beth Levin's classification of English verbs and then VerbNet (Kipper et al., 2000), its realisation as an on-line computational linguistic resource. Then, I show how to extract paraphrastic patterns for verbs from these resources and how to extend these patterns by using the lexical semantic knowledge encoded in WordNet.

3.1 Levin's Verb Classes

Verbs are one of the most central syntactic categories in language and have been studied at length by linguists. One of the most prominent approaches to verbs is that of Levin (1993) who proposes a theory of the verb lexicon. Starting from the assumption that lexical knowledge is more than knowledge of idiosyncratic word-specific properties, and guided by the intuition that similar syntactic behaviour is a clue of similar semantic properties, Levin (1993) proposed a classification of English verbs based on their alternations.

3.1.1 Alternations as Semantic Tests

Starting from the intuition that the syntactic realisation of the arguments of a verb (i.e. their grammatical function and syntactic type) is predictable from the meaning of that verb, Levin (1993) describes the semantics of verbs by considering the set of alternations each verb (dis)allows.

As we saw in the previous chapter, *alternations* are syntactic variations conveying the same (or a similar) lexical meaning. More specifically, alternations are defined by Levin as changes in the expressions of verb arguments which are sometimes accompanied by changes of meaning. For example, the sentences in (39) are examples of *Instrument Subject Alternation* which relates two different syntactic constructions allowed for the verb *open* which are paraphrastic: a first construction in which the grammatical subject (*this key*) of the verb fulfils the *instrument* role of the verb and a second construction in which the *instrument* role is realised as a PP (*with this key*).

(39) a. This key opens the safe. *Instrument Subject Alternation*

b. The safe opens with this key.

In order to build the classification of verbs, Levin (1993) considered solely *intracategorial alternations* which she also calls *diathesis alternations*, i.e. syntactic variations which present a different linking between grammatical functions and thematic roles of the arguments of the verbs. She identifies 79 such verb alternations. It is worth noting that not all diathesis alternations analysed by Levin are content preserving alternations. Consider, for example, the set of transitivity alternations shown in Figure 3.1. These alternations involve a change in the transitivity of a verb.

Transitivity alternations are often accompanied by a change in meaning, as the intransitive variant lacks an argument. In (40) for instance, the subject of the transitive variant is unexpressed in the intransitive variant (40b), while in (41) it is the object that remains unexpressed. On the contrary, the *Understood body-part Object Alternation* in (42) is content preserving as the variants (42a) and (42a) are paraphrastic due to the fact that the verb *floss* already contains as part of its meaning the unexpressed object *teeth*.

(40) a. The butcher cuts the meat. *Middle Alternation*

b. The meat cuts easily.

(41) a. Mike ate the cake. *Underspecified Object Alternation*

b. Mike ate.

1. **Transitivity Alternations**
    (a) Object of Transitive = Subject of Intransitive Alternation
        i. Middle Alternation
           The butcher cuts the meat./ The meat cuts easily.
        ii. Causative Alternations
            A. Causative/Inchoative Alternation
               Janet broke the cup./ The cup broke.
            B. Induced Action Alternation
               Sylvia jumped the horse over the fence./ The horse jumped over the fence.
            C. Other Instances of Causative Alternations
               The visitor rang the bell./ The bell rang.
        iii. Substance/Source
             Heat radiates from the sun. / The sun radiates heat.
    (b) Unexpressed Object Alternation
        i. Unspecified Object Alternation
           Mike ate the cake./ Mike ate.
        ii. Understood Body-Part Object Alternation
            I flossed my teeth./ I flossed
        iii. Understood Reflexive Object Alternation
             Jill dressed herself hurriedly./ Jill dressed hurriedly.
        iv. Understood Reciprocal Object Alternation
            Anne met Cathy./ Anne and Cathy met.
        v. PRO-arb Object Alternation
           The sign warned us against skating on the pond./ The sign warned against skating on the pond.
        vi. Characteristic Property Alternations
            A. Characteristic Property of Agent Alternation
               The dog bites people./ The dog bites.
            B. Characteristic Property of Instrument Alternation
               I cut the bread with this knife./ This knife cut the bread.
        vii. *Way* Object Alternation
             They pushed their way through the crowd./ They pushed through the crowd.
        viii. Instructional Imperative
              Bake the cake for 30 minutes./ Bake for 30 minutes.
    (c) Conative
        Paula hit the fence./ Paula hit at the fence.

Figure 3.1: Levin's Transitivity Alternations

(42) a. *I flossed my teeth.* *Understood body-part Object Alternation*

b. *I flossed.*

Now, I will illustrate with an example how verbs can be classified based on the type of alternations in which they participate.
Consider the four verbs of contact *cut, break, touch, hit*. As sentences (43) show, they are all transitive; they all can have an object NP.

(43) a. *Margaret cut the bread.*

b. *Janet broke the vase.*

c. *Terry touched the cat.*

d. *Carla hit the door.*

But they behave differently with respect to the following three alternations.

Middle alternation As the examples in (44) show, only the verbs *cut* and *break* can participate in this alternation.

(44) a. *The bread cuts easily.*

b. *Glass vases break easily.*

c. **Cats touch easily.*

d. **Doors hit easily.*

Conative alternation As the examples in (45) show, only the verbs *cut* and *hit* present this syntactic frame.

(45) a. *Margaret cut at the bread.*

b. **Janet broke at the vase.*

c. **Terry touched at the cat.*

d. *Carla hit at the door.*

Body-part possessor alternation As the examples in (46) show, by considering this alternation it is possible to distinguish between the behaviour of the verb *break*, for which this construction is not allowed, and the other three verbs in the semantic class of the verbs of contact.

(46) a. *Margaret cut Bill on the arm. (cf. Margaret cut Bill's arm)*

b. **Janet broke Bill at the finger. (cf. Janet broke Bill's finger)*

c. *Terry touched Bill on the shoulder. (cf. Terry touched Bill's shoulder)*

d. *Carla hit Bill on the back. (cf. Carla hit Bill's back)*

The different behaviour of the verbs of contact with respect to these three alternations is summarised in Figure 3.2. It suggests that the class of verbs of contact is not a homogeneous semantic class and should be subcategorised into four subclasses.

Alternation	Touch	Hit	Cut	Break
conative	-	+	+	-
body-part possessor ascension	+	+	+	-
middle	-	-	+	+

Figure 3.2: Levin’s Verb-class: Hit Verbs

The behaviour shown by these four verbs is common to other verbs of contact so that the following four subclasses can be defined:

- **Touch-Verbs**: kiss, sting, tickle
- **Hit-Verbs**: bash, hammer, tap
- **Cut-Verbs**: chip, hack, scratch
- **Break-Verbs**: hack, split, tear

3.1.2 Example of Levin’s Verb Classes

A Levin verb class is fully individuated by:

- the set of verbs that are members of the class,
- the set of alternations common to the verbs in the class
- the core semantics shared by the members of the class

An example will show how this information is represented by Levin. Figure 3.3 describes the *Hit-Verbs* class, a subclass of the verbs of contact.
Each verb class lists the alternations possible for all members of the class. The members of the *Hit-Verbs* class, for instance, can participate in the *With/Against* Alternation, in the *Conative* Alternation, in the *Body-Part Possessor Ascension*, in the *Together Reciprocal* and in the *Instrument Subject* Alternations.
A verb class also contains a description of the semantics shared by the members of the class. The members of a verb class are generally not synonymic but they share some core semantics which are recorded in the representation of the class. In the case of the *Hit-Verbs* these semantics correspond to the discursive description in Figure 3.3. Thus, for example both verbs *hit* and *kick* have a core semantic component describing moving an entity in order to bring it in contact with a second entity. But they are not synonyms. The verb *hit*, in fact, describes the general action of *hitting* while the verb *kick* specifies the hitting action in *hitting with the foot*.

Class Members

bang, bash, batter, beat, bump, butt, dash, drum, hammer, hit, kick, knock, lash, pound, rap, slap, smack, smash, strike, tamp, tap, thump, thwack, whack

Alternations

1. *With/Against* Alternation
   a. Paula hit the stick against/on the fence.
   b. Paula hit the fence with the stick.

2. *Conative* Alternation
   a. Paula hit the fence (with the stick).
   b. Paula hit at the fence (with the stick).

3. *Body-Part Possessor Ascension* Alternation
   a. Paula hit Deirdre on the back.
   b. Paula hit Deirdre's back.

4. *Together Reciprocal* Alternation
   a. Paula hit one stick against another.
   b. Paula hit the sticks together.

5. *Instrument Subject* Alternation
   a. Paula hit the fence with the stick.
   b. The stick hit the fence.

Semantics

These verbs describe moving one entity in order to bring it into contact with another entity, but they do not necessarily entail that this contact has any effect on the second entity.

Figure 3.3: Levin's Verb-class: Hit Verbs

3.2 Verbnet: Extending Levin's Verb Classes

Levin classes, although an important theoretical starting point, do not provide information that is precise enough to be used to build a computational verb lexicon. Levin classes have two fundamental limitations that makes their utility as general classification schemes problematic.

Semantic ambiguity of verbs Many verbs are listed by Levin in more than one class. Often these classes present different and conflicting sets of syntactic frames. For example, the verbs *push, pull, tug, shove, kick* are listed both in the *carry* class and in the *push/pull* class. But while the members of the first class cannot participate in the conative alternation (see sentence (47)), the members of the second class can (see sentence (48)).

(47) a. **The mother carried at the baby.*

(48) a. *John kicked at the ball.*

In other words, since *push, pull, tug, shove, kick* are members of both classes, they are predicted to be both compatible and incompatible with the conative alternation. To explain this contradiction, it is necessary to hypothetise distinct meanings for these verbs and to assign verb meanings (rather than verb forms) to verb classes.

Lack of semantic representation Levin does not assign any systematic semantic representation to verbs in the verb classes (see as an example the description of the semantic of the *Hit* verbs given in Figure 3.3) and in particular there is no explicit link between semantic roles and syntactic arguments of verbs.

In order to solve both these problems, VerbNet (Kipper et al., 2000), a broad-coverage domain-independent computational verb-lexicon, was developed.
VerbNet encodes syntactic and semantic information for about 4000 English verbs. The verbs are organised in classes which refine Levin's classes and capture generalisations about the regular association between syntactic and semantic verb properties.
The problem of semantic ambiguity of verbs is solved in VerbNet first, by linking each verbal item in a class to a WordNet sense, in this way it is possible to assign different senses of a verb which exhibit different syntactic properties to different classes, and second, by refining Levin's classes in *intersective classes* (Dang et al., 2000). The idea behind intersective classes is that if the intersection between two or more Levin classes is not empty, the class individuated by the members in the intersection defines a more fine grained class which exhibits more coherent sets of syntactic frames and associated semantic components (Dang et al., 1998). Further in VerbNet, syntactic frames containing prepositional phrases or adverbs providing a regular extension of meaning to the core sense of many verbs are also considered.

The second problem of the Levin verb classification, the lack of a systematic semantic representation for verbs, is solved in VerbNet by assigning a flat semantic representation, represented as a conjunction of predicates, to each syntactic frame of a class.

1.	**Actor**	used in verb of communication classes when both arguments are symmetrical
2.	**Agent**	a human or an animate subject
3.	**Asset**	used with ’currency’ as selectional restriction
4.	**Attribute**	refers to a quality of something that is being changed, used with ’scalar’ as selectional restriction
5.	**Beneficiary**	the entity that benefits from some action
6.	**Cause**	the cause of some event
7.	**Location**	underspecified destination, source or place
8.	**Destination**	end point of the motion, or direction towards which the motion is directed
9.	**Source**	start point of the motion
10.	**Experiencer**	used for a participant that is aware or experiencing something
11.	**Extent**	used to specify the range or degree of change
12.	**Instrument**	used for objects or forces that come into contact with an object and cause some change in them
13.	**Material**	start point of transformation
14.	**Product**	end result of transformation
15.	**Patient**	used for participant that is undergoing a process or that has been affected in some way
16.	**Predicate**	used for classes with a predicative complement
17.	**Recipient**	target of the transfer
18.	**Stimulus**	used by verbs of perception for events or objects that elicit some response from an Experiencer
19.	**Theme**	used for participants in a location or undergoing a change of location
20.	**Time**	used to express time
21.	**Topic**	topic of communication verbs

Figure 3.4: VerbNet Semantic Roles

Moreover, each syntactic argument of a frame is associated with a thematic role. VerbNet uses the set of 21 thematic roles[8] shown in Figure 3.4.
More specifically, a VerbNet class has the following components:

- The set of English verbs belonging to that class, each verb being annotated with the WordNet meaning(s) relevant to that class
- The set of theta roles which can be mapped to the arguments of these verbs
- Selectional restrictions on the arguments
- A set of frames consisting of an identifier, an example, a syntactic description and decompositional semantics common to all verbs in that class
- The set of superclasses of that class if any exists. The frames of these upper classes are then inherited.

As an example, consider Figure 3.5 picturing the VerbNet frame for the *meander-47.7* class. Members of this class are verb senses such as *crawl*, *cut*, *drop, go*. The thematic roles used to describe the semantics of each syntactic pattern defined for the class are the *Theme* and the *Location* roles with their selectional restrictions, thus for instance in this class only a constituent with a semantic feature `+concrete` can fill the *Location* role.
The syntax-semantic interface for the verbs in this class is represented with a set of frames described by:

- an example sentence
- the syntactic pattern realised in the example
- an event semantic representation of the semantics of the pattern

For instance, the *Locative Inversion* frame is described by the example sentence

(49) a. Through the valley meanders the river.

and is associated with the syntactic frame `Prep[+path] Location V Theme` describing a `PP V NP` syntactic configuration in which the preposition in the PP is constrained to be a locative preposition, i.e. a path ($[+path]$) denoting preposition and the subject NP is assigned a *Theme* role. The associated semantic formula

`Prep(during(E),Theme,Location)` $\wedge$ `exist(during(E),Theme)`

describes an event *E* during which the *Theme* object, which is supposed to exist, moves *to, in, through* (i.e. a preposition describing a path) a location *Location*.

[8]Kipper-Schuler (2005) claims that although it is difficult to define an exhaustive set of theta roles which can describe all arguments of verbs, the 21 roles used in VerbNet were chosen so to include the most common used roles and to allow a better specification of the semantics of the verb classes still permitting generalisations about verb behaviour.

Members	
	cascade(1), climb(4), crawl(1), cut, drop(1,2), go(7), meander(1), plunge(1), run(3), straggle(2), stretch(1), sweep(5), tumble, turn, twist(1), wander(4), weave(4), wind(1 2)
Thematic Roles and Selectional restrictions	Location[+concrete] Theme[+elongated]
Frames	
	Intransitive (+ path PP) "The river runs through the valley" `Theme V Prep[+path] Location` Prep(during(E),Theme,Location) exist(during(E),Theme)
	Locative Inversion "Through the valley meanders the river" `Prep[+path] Location V Theme` Prep(during(E),Theme,Location) exist(during(E),Theme)
	PP-NP Expletive-There Subject "There meanders through the valley a river" `there V Prep[+path] Location Theme` Prep(during(E),Theme,Location) exist(during(E),Theme)
	NP-PP Expletive-There Subject "There meanders a river through the valley" `there V Theme Prep[+path] Location` Prep(during(E),Theme,Location) exist(during(E),Theme)

Figure 3.5: VerbNet representation of the *meander-47.7* class

3.3 Extracting Verb Paraphrastic Patterns from VerbNet

For the treatment of paraphrases, the information contained in VerbNet is useful for several reasons. First, VerbNet documents the alternations of each verb in a verb class. For instance, the *meander-47.7* class allows the following meaning-preserving alternations: *Intransitive + Path PP*, *Locative Inversion*, *PP-NP Expletive-There Subject* and *NP-PP Expletive-There Subject*.
This information permits the automatic extraction of the set of generic paraphrastic structures or paraphrastic rules allowed for each verb in a class.
Therefore, for each verb in the *meander-47.7* class, the following paraphrastic constructions can be predicted:

1. $\texttt{NP}_{1}\ \texttt{V}_{meander}\ [\texttt{Prep NP}_{2}]_{PP}$ *Intransitive + Path PP*
2. $[\texttt{Prep NP}_{2}]_{PP}\ \texttt{V}_{meander}\ \texttt{NP}_{1}$ *Locative Inversion*
3. $\texttt{There V}_{meander}\ [\texttt{Prep NP}_{2}]_{PP}\ \texttt{NP}_{1}$ *PP-NP Expletive-There Subj.*
4. $\texttt{There V}_{meander}\ \texttt{NP}_{1}\ [\texttt{Prep NP}_{2}]_{PP}$ *NP-PP Expletive-There Subj.*

Further, as VerbNet associates with each verb class a thematic grid and decompositional semantics, it becomes possible to develop a parser which, based on this knowledge (knowledge of the alternations of a verb and of its semantic representation) can build identical semantic representations for alternations paraphrases.

(50) a. The river meanders through the valley.

b. Through the valley meanders the river.

c. There meanders through the valley a river.

d. There meanders a river through the valley.

Thus, for example, using the thematic role information associated in VerbNet with the *meander-47.7* class, all sentences in (50) can be assigned the same basic semantic representation :

$$\texttt{River(R)} \wedge \texttt{Valley(V)} \wedge \texttt{Meander(E)} \wedge \texttt{Location(E,V)} \wedge \texttt{Theme(E,R)}$$

For paraphrase recognition this is enough; for deeper semantic treatment involving inference for instance, decompositional semantics might also be useful.

3.4 Extending Verb Paraphrastic Patterns with WordNet

Another feature of VerbNet which makes it attractive for the treatment of paraphrases is its linking with WordNet.

WordNet (Miller et al., 1993; Fellbaum, 1998) is a semantic lexicon developed for the English language by the Cognitive System Laboratory at Princeton University. In this lexicon words are grouped in synsets, i.e. sets of lexical items which lexicalise the same concepts, and thus are synonyms. Different senses of a word correspond to different synsets. WordNet synsets are connected through semantic relations which vary depending on the type of the synset, (e.g. *hyponymy, hyperonymy* for nouns *troponymy, hyponymy* for verbs, etc.).
WordNet (Fellbaum, 1990) organises verbs in 11,500 synsets, which also include verb collocations. The semantic relations used to structure verbs are hyperonymy, entailment and troponymy.
Given that X and Y are two different verbs, these relations are defined as given below.

Hyperonymy Y is a hyperonym of X
: if *to X* is a kind of *to Y*. For example, the verb *move* is a hyperonym of the verb *meander*, as *meander* is a kind of moving, namely it means *move in a sinuous course*.

Entailment X entails Y
: if by doing X someone must be doing Y. For example, the verb *snore* entails the verb *sleep* as if someone is *snoring* he is also *sleeping*; nobody can snore if he is not sleeping.

Troponymy X is a troponym of Y
: if *to X* is equal to *to Y in some manner*. For example, the verb *poison* is a troponym of the verb *kill* as *to poison* means *to kill in some manner*, namely by poisoning. Troponymy is thus the verb specific relation which corresponds to the hyponymy relation defined for nouns.

Figure 3.6 shows the WordNet entry for the verb *meander*.
The fact that each verb in VerbNet is linked to a WordNet synset permits (i) to extend VerbNet classes by enriching the set of verbs in a class with their synonyms[9] and (ii) to structure VerbNet classes by considering the lexical relations (e.g. troponymy, antonymy, etc.) which hold between their members and thus (iii) to extend the set of paraphrastic patterns which can be extracted from VerbNet by allowing for a given verb

[9] Verbs, members of the same synset, are weak synonyms, i.e. they cannot always be substituted in a given register or context. Consider, for example, the different register selected when using the verb *begin* and the verb *commence*, or the different selectional restrictions associated with the usage of the verbs *rise* and *ascend* (e.g. a temperature can rise but a temperature “ascending” is much less likely). These differences do not affect paraphrase recognition as implemented in this thesis, thus all the members of a WordNet synset can be considered members of the same Verbnet class.

Sense 1	
meander	{weave, wind, thread, meander, wander} (to move or cause to move in a sinuous, spiral or circular course; "the river winds through the hills"; "the path meanders through the vineyards")
	Semantic Relations:
	Hyperonyms $\Longrightarrow$ { travel, go, move, locomote} *Troponyms* $\Longrightarrow$ { snake}

Figure 3.6: WordNet Representation of the Verb *meander*

to multiply the paraphrastic rules found through content preserving alternations with the number of synonyms of that verb:

$$|\text{Paraphrases}| = |\text{Paraphrastic Patterns}| \times |\text{Synset(Verb)}|$$

Figure 3.7 shows, as an example, the result of extending the VerbNet *meander-47.7* class with the synsets extracted from WordNet for all its members.

Members	
	$Syn_{meander(1)}$: *{meander(1), thread(1), wander(4), weave(4), wind(1 2)}*
	$Syn_{go(7)}$: *{go(7), run(3)}*
	$Syn_{cascade(1)}$: *{cascade(1), cascade down(1)}*
	$Syn_{crawl(1)}$: *{crawl(1), creep(1)}*
	$Syn_{straggle(2)}$: *{straggle(2), sprawl(2)}*
	$Syn_{stretch(1)}$: *{stretch(1), stretch down(1)}*
	$Syn_{sweep(5)}$: *{sweep(5), span(1), traverse(2), cross(5)}*
	...

Figure 3.7: Adding structure to VerbNet: *meander-47.7* class

In the case of the verb *meander*, for instance, the set of synonyms retrieved from WordNet for sense 1 of *meander* is *weave, wind, thread, meander, wander*. By combining this information with the knowledge of alternations given by VerbNet, the following 20 paraphrastic sentences can be predicted (|paraphrastic patters| $\times$ |synonyms| = 4$\times$5).

(51) *1. The river meanders through the valley*
2. Through the valley meanders the river
3. There meanders through the valley a river
4. There meanders a river through the valley
5. The river weaves through the valley

6. Through the valley weaves the river
7. There weaves through the valley a river
8. There weaves a river through the valley
9. The river winds through the valley
10. Through the valley winds the river
11. There winds through the valley a river
12. There winds a river through the valley
13. The river threads through the valley
14. Through the valley threads the river
15. There threads through the valley a river
16. There threads a river through the valley
17. The river wanders through the valley
18. Through the valley wanders the river
19. There wanders through the valley a river
20. There wanders a river through the valley

Members	
	despise(1), detest(1), disdain(1), dislike(1), enjoy(1 3 5), fear(2 4 5), hate(1), like(2 3), love(1 2 3), regret(1 2 5)
Frames	
	TO-INF-SC "I love to write" Experiencer V Theme
	ING-SC/BE$_{ing}$ "I love writing" Experiencer V Theme
	...

Figure 3.8: VerbNet: *admire-31.2-1* class

By considering other lexical semantic relations such as antonymy, hyperonymy and troponymy, it is possible first to further extend the set of paraphrastic patterns found in VerbNet (by hand of antonymy) and second, to also handle with verb entailment (by using hyperonymy and troponymy).
As an example, consider the VerbNet class *admire-31.2-1* shown in Figure 3.8. By considering the lexical knowledge encoded in WordNet about synonymy and antonymy the class can be structured as shown in Figure 3.9, thus allowing the recognition of the paraphrastic patterns in (52) which capture the antonymic relations which hold between the members of the verb class.

(52) *1. John does not like fishing.*
2. John does not like to fish.

Members

Syn$_1$: {*like(2 3)*}
→ `antonyms:` *dislike(1)*
Syn$_2$: {*love(1 2 3), enjoy(1 3 5)*}
Syn$_3$: {*dislike(1)*}
→ `antonyms:` *like(2 3)*
Syn$_4$: {*hate(1), detest(1)*}
Syn$_5$: {*despise(1), disdain(1)*}
Syn$_6$: {*regret(1 2 5)*}
Syn$_7$: {*fear(2 4 5)*}

Figure 3.9: Adding structure to VerbNet: *admire-31.2-1* class

3. *John dislikes fishing.*
4. *John dislikes to fish.*

Because the VerbNet classes are often semantically homogeneous classes, the approach proposed here provides a handle on entailment between verbs. As in Levin's work, the VerbNet grouping of verbs into classes is based on syntactic criteria: verbs sharing the same set of alternations are grouped together. However, the driving idea is that shared syntactic properties reflect shared semantic properties so that in effect, verbs belonging to the same VerbNet class are likely to be semantically similar. This is particularly clear in the *admire-31.2-1* class presented above where the verbs of that class are very close ontologically: they all express a psychological state[10].

Now, it is possible to use the WordNet hierarchy to order the set of verbs included in a VerbNet class with respect to hyperonymy and troponymy. By using the concept hierarchy thus defined, entailment between verbs can then be checked.

Thus, for example the VerbNet class *admire-31.2-1*, by considering the lexical knowledge encoded in WordNet about troponymy, can be structured as shown in Figure 3.10 thus allowing the recognition of the entailments[11] in (54).

(54)

[10]Although, these verbs can have different implications.

[11]It is worth stressing that the information about antonymy provided by WordNet is not sufficient to recognise entailment. WordNet in fact makes no distinction between binary and contrary oppositions. Thus, in the case of the verbs like *like* and *dislike* (which are contraries) this can yield to the recognition of false entailments as in (53).

(53) *John does not dislike cooking.* → *John likes cooking*

Members

Syn$_1$: {*like(2 3)*}
$\rightarrow$ `antonyms:` *dislike(1)*
$\rightarrow$ `direct troponyms:` {*love(1 2 3), enjoy(1 3 5)*}

Syn$_2$: {*love(1 2 3), enjoy(1 3 5)*}

Syn$_3$: {*dislike(1)*}
$\rightarrow$ `antonyms:` *like(2 3)*
$\rightarrow$ `direct troponyms:` {*hate(1), detest(1)*}

Syn$_4$: {*hate(1), detest(1)*}
$\rightarrow$ `direct troponyms:` {*despise(1), disdain(1)*}

Syn$_5$: {*despise(1), disdain(1)*}

Syn$_6$: {*regret(1 2 5)*}
$\rightarrow$ `direct troponyms:` {*fear(2 4 5)*}

Syn$_7$: {*fear(2 4 5)*}

Frames

TO-INF-SC
"I love to write"
Experiencer V Theme

ING-SC/BE$_{ing}$
"I love writing"
Experiencer V Theme

...

Figure 3.10: Adding structure to VerbNet: *admire-31.2-1* class

1.	*John dislikes cooking.*	$\rightarrow$	*John does not like cooking.*
2.	*John dislikes cooking.*	$\rightarrow$	*John does not love cooking.*
3.	*John dislikes cooking.*	$\rightarrow$	*John does not enjoy cooking.*
4.	*John dislikes cooking.*	$\rightarrow$	*John does not like to cook.*
3.	*John detests/hates cooking.*	$\rightarrow$	*John dislikes cooking.*
4.	*John detests/hates cooking.*	$\rightarrow$	*John dislikes to cook.*
5.	*John detests/hates cooking.*	$\rightarrow$	*John does not love cooking.*
6.	*John detests/hates cooking.*	$\rightarrow$	*John does not enjoy to cook.*
7.	*John disdains/despises cooking.*	$\rightarrow$	*John detests cooking.*
8.	*John disdains/despises cooking.*	$\rightarrow$	*John hates to cook.*
9.	*John disdains/despises cooking.*	$\rightarrow$	*John dislikes cooking.*
10.	*John disdains/despises cooking.*	$\rightarrow$	*John dislikes to cook.*
11.	*John disdains/despises cooking.*	$\rightarrow$	*John does not like cooking.*
12.	*John disdains/despises cooking.*	$\rightarrow$	*John does not enjoy to cook.*
13.	*John likes cooking.*	$\rightarrow$	*John does not dislike cooking.*
14.	*John likes cooking.*	$\rightarrow$	*John does not detest to cook.*
15.	*John loves/enjoys cooking.*	$\rightarrow$	*John likes cooking.*
16.	*John loves/enjoys cooking.*	$\rightarrow$	*John likes to cook.*
17.	*John loves/enjoys cooking.*	$\rightarrow$	*John does not dislike cooking.*
18.	*John loves/enjoys cooking.*	$\rightarrow$	*John does not detest to cook.*
19.	*John does not like cooking.*	$\rightarrow$	*John dislike/detest cooking.*
20.	*John does not like cooking.*	$\rightarrow$	*John dislike/detest to cook.*

3.5 Summary

In this chapter, I presented Levin's classification of English verbs. I then described VerbNet (Kipper et al., 2000), an on-line lexical resource which implements Levin's ideas and I showed how to use such a resource to systematically acquire paraphrastic patterns for verbs and how to extend them by also considering the lexical knowledge encoded in WordNet(Fellbaum, 1998).
In the next chapter, I present a symbolic approach to verb-paraphrasing which uses the resources presented in this chapter.

Computational Treatment of Verb Paraphrases: a Linguistic-Based Approach

In this chapter building on (Amoia and Gardent, 2005), I present a symbolic treatment of paraphrases which focuses on verb alternations and lexical synonymy. For this type of paraphrases, I present a robust system which assigns two such paraphrases one and the same semantic representation.
Robustness is achieved using a cascaded finite state parser (Xerox Incremental Parser henceforth, XIP) based on layered grammars (grammars are ordered and the rules in each grammar can refer to the representation produced by the preceding grammars). The XIP system is robust in that it always delivers a single output although the parse produced may be partial and underspecified in case a full analysis cannot be performed. The coverage achieved by the extended XIP system (the system extended with a grammar layer for the recognition of alternation based paraphrases) reflects the knowledge encoded in the XIP grammars, in VerbNet and in WordNet. That is, it integrates the detailed and extensive knowledge of syntax, alternations and lexical relations encoded in these three resources. As no benchmark for alternation based paraphrase recognition was available however, the system evaluation was restricted to a hand built benchmark of roughly 900 sentence pairs extracted from VerbNet.

The chapter is structured as follows: I start by presenting the XIP parser and the type of representations produced with some examples. I go on to show how to extend XIP to integrate VerbNet and WordNet information so as to assign paraphrases the same semantic representation. Finally, I present an evaluation of the system based on a set of annotated examples extracted from VerbNet.

4.1 XIP

Robustness, that is, the ability to process real world textual data, is an important desideratum of natural language processing systems both from a theoretical and from a practical point of view. Theoretically it is important, because testing theories with non artificial data is a requirement of the scientific enterprise and practically, because it is necessary in real world applications. As accurately summarised in (Ait-Mokhtar et al., 2002), three main types of approaches to robustness can be distinguished: those based on deep grammars and using special mechanisms in order to recover an analysis when parsing fails or to rank analyses in case of overgeneration; those based on probabilistic parsing approaches and those adopting a shallow approach to parsing usually based on finite state techniques.

The XIP parser belongs to the third type of approach and guarantees robustness by adopting incrementality: the input sequence is processed by a layered grammar, each grammar layer being applied sequentially. As the input is processed, it is either enriched or left unchanged – the output is the input sequence as annotated by the sequential application of the rules from the different layers. By ordering the grammar rules appropriately, data which is either infrequent or incorrect (e.g., sentences violating verb/subject agreement) can therefore be handled. It suffices to place the rules handling that data last. Since the data does not conform to the rules governing the most frequent data (which are placed early in the grammar layers), these rules do not take effect and the rules governing the infrequent or incorrect data can apply.

Based on finite state techniques, the parser is also reasonably efficient running at a speed of 1,300 words per second on a Pentium II 50 where processing time includes tokenisation, morphological analysis, part of speech disambiguation, chunking and dependency parsing.

I use XIP version 3.10 (2000-2001) as developed at Xerox Research Europe. This version includes the NTM tokeniser and morphological analyser based on finite states technology, the HMM statistical POS tagger and a grammar for English which includes two types of subgrammars, namely, chunking and dependency grammars.

Figure 4.1 shows the output of the original version of XIP after parsing of sentence (55).

(55) *Brutus killed Caesar.*

The **tokeniser** segments the input strings into tokens i.e., words, punctuation signs and multi-word expressions whilst the **morphological analyser** assigns a part of speech to each token based on regular expressions over words forms.

The **chunking grammar** describes constituency structure and consists of layered groups of chunking rules which are either ID (immediate dominance) rules applying to partially ordered bags of nodes or LP (linear precedence) rules applying to ordered subsequences.

Chunking rules are grouped into layers, each layer applying to the output of the preceding layers, thus allowing for the production of chunking trees with reasonable depth. Within one layer, only the first applicable rule (if any) is applied.

Figure 4.2, for example, shows five sequence rules organised through three layers. The rules in the first layer (1 $>$) build an NP chunk if the list of nodes contained in the current XIP stack contains (i) a determiner which is followed by a noun, (ii) a determiner which is followed by a noun possibly preceded by an adjective or (iii) a pronoun. The

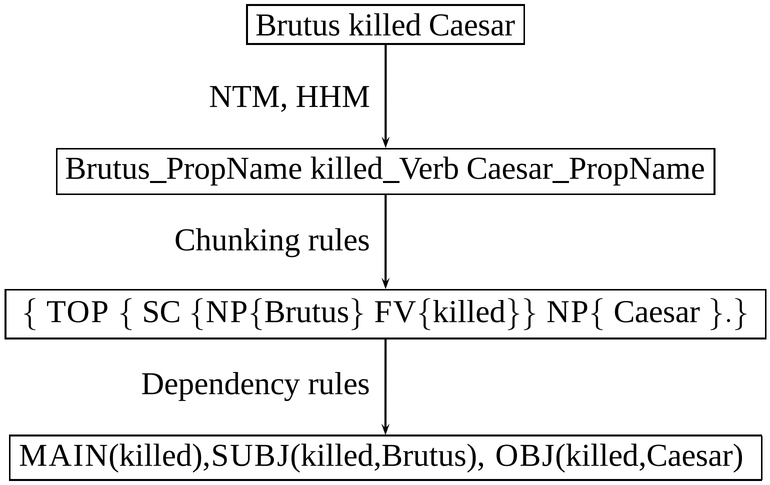


Figure 4.1: XIP representation of the sentence *Brutus killed Caesar.*

rule in the second layer ($2 >$) builds a VP chunk if the XIP stack contains a verb possibly preceded by an adverb, while the rule in the third layer ($3 >$) builds a SC (sentence clause) chunk if a NP and a VP chunk were already recognised.
The chunking rules are applied in the ordered sequence: layer N-1 before layer N, so that the NP chunk is built before the VP chunk and SC chunk can be built first after the NP and VP chunks. In a given layer, the first rule which matches the input (scanned from left to right) is applied.
For instance, by scanning the input sentence (55) the first rule of layer 1 which applies is $NP = PropName$. Thus, after processing the first layer the stack contains two NP nodes, i.e. Brutus_NP and Caesar_NP. Thereafter, the input is scanned again for rules in

$1 >$	NP = Det, Noun.
$1 >$	NP = Det, Adj*, Noun.
$1 >$	NP = Pron.
$1 >$	NP = PropName.
$2 >$	VP = adv*, Verb.
$3 >$	SC = NP, VP.

Figure 4.2: Example Xip chunking rules

the second layer to match. The information in the stack bound to the verb *killed* matches the rule $VP = adv^*, Verb$ and the stack is updated with the killed_VP chunk. Finally, layer 3 applies to the stack and yields the chunked tree shown in Figure 4.3, where the input sentence is chunked as a SC followed by an NP[12]

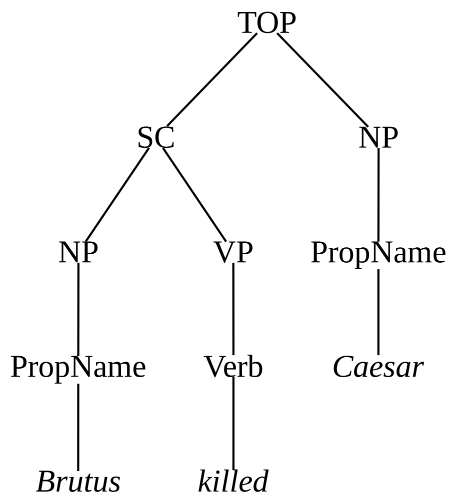


Figure 4.3: XIP chunked tree for *Brutus killed Caesar.*

The **dependency grammar** supports the specification of (functional, thematic, semantic, anaphoric, etc.) relations between words or chunks and is based on the (layered) specification of groups of dependency rules of the form:

```
|<subtree_pattern>|
   if <condition>
   <dependency_term1> <dependency_term2>...<dependency_termN>
```

- `subtree_pattern` is a tree matching expression describing structural properties of part of the input tree. Such expression states a constraint on the form of the chunked sentence/text. The `subtree_pattern` expression in (56), for example, describes the syntactic structure of sentences like (59) in which an NP chunk (i.e. $NP\{?*, \#1[last : +]\}$) is immediately followed by a VP chunk (i.e. $VP\{?*, \#2[last : +]\}$). If this pattern expression matches part of the input tree, the heads of the NP and of the VP are bound respectively to the variable $\#1$ and $\#2$ which can be referred to in the `condition`. The `subtree_pattern` slot of a rule can also be empty.

  (56) $|NP\{?*, \#1[last : +]\}, VP\{?*, \#2[last : +]\}|$[13]

[12]For each chunked sentence, a root node is created by the chunker and linked to all the longest chunks found.

[13]Meaning of XIP notation:

- `condition` is any Boolean expression built up from dependency terms, linear order statements and the conjunction (i.e. &), disjunction (i.e. $||$) and negation (i.e. $\sim$) operators. Such an expression states a constraint on previously computed dependency relations. The `condition` expression in (57), for example, imposes the restriction that the subject should not occur in a passive clause, i.e. if there exists a (previously computed) SUBJ dependency relation between words $\#1$ and $\#2$ then it should not have the feature *passive* in order for the current dependency rule to apply.

  (57) **if** $\sim$SUBJ*[passive:+]*$(\#1, \#2)$

- `dependency_term` is a term of the form `name`[`flist`]`(f1, ..., fN)`, describing a dependency relation with name `name` and arguments `f1, ..., fN`. The dependency relation can be associated with a list of features [`flist`]. The `dependency_term` in (58), for example, defines a SUBJ dependency relation between the words at nodes $\#1$ and $\#2$ and an OBJ dependency relation between the words at nodes $\#3$ and $\#2$. The OBJ dependency relation is associated with a list of features $[pro : +]$ which further specifies that the object of the verb should be a pronoun.

  (58) SUBJ*(#1,#2),* OBJ*[pro:+](#3,#2)*

A dependency rule should be read as follows: if the chunked sentence meets the constraint imposed by the subtree pattern defined in the rule, and if the previously computed dependency relations meet the constraint stated in the conditions part of the rule, then the dependency relations expressed through the dependency terms can be stated.
Figure 4.4 shows dependency rules which capture grammatical relations such as those realised in sentence (59).

(59) *John always enjoys good wine.*
SC [[*NP John*][*VP always enjoys*]][*NP good wine*]]

Dependency rules are applied in cascade and rules from a given layer can apply to the output of a previous layer. That is, they apply on both the initial constituent tree and the incremental set of previously computed dependencies.

- | ... | defines a set of adjacent trees
- Cat{ ... } defines a tree with root category Cat and structure { ... }
- the comma (,) separates two adjacent constituent nodes
- $?*$ marks an arbitrary constituent
- $\#N$ marks a node with a node variable N
- $\#N[feat_1 : +, ..., feat_N : -]$ marks a node with a node variable N and a features list $[feat_1 : +, ..., feat_N : -]$
- $\#$ marks a node with an anonymous node variable
- the features *last* marks the last element of the chunk.

1: $| NP\{?*, \#1[last:+]\}, VP\{?*, \#2[last:+]\} |$
SUBJ$(\#1, \#2)$

2: $| SC\{NP\{?*, \#1[last:+]\}, VP\{?*, \#2[last:+]\}\}, NP\{?*, \#3[last]\} |$
if ~(SUBJ(#,#2))
SUBJ$(\#1, \#2)$, OBJ$(\#3, \#2)$

3: $| SC\{?*, FV[trans:+]\{?*, \#1[last:+]\}\}, NP[time:\sim]\{?*, \#2[last:+]\} |$
VCOMP$[dir:+](\#1, \#2)$

Figure 4.4: Example of Xip Dependency Rules

For instance, given the rule ordering given in Figure 4.4 and sentence (59) as input, the first rule which applies is rule 1. It creates a SUBJ (subject) dependency between the word *John* and the word *enjoys* because the chunk tree assigned to this sentence by the chunking rules has the form defined in the rule, namely an NP chunk followed by a VP chunk and there is no additional constraints stated by the rule. Thus, the SUBJ dependency relation can be defined between the words at nodes $\#1$ and $\#2$: SUBJ$(\#1, \#2)$ namely *John* and *enjoys*.

Rule 2 places a constraint on previously computed dependencies, namely that there is no SUBJ dependency already assigned between the VP and some other chunk.

The more complex rule 3 recognises a VCOMP (verb-complement) dependency such as that between *wine* and *enjoys* in sentence (59). This rule constrains the FV chunk to be such that a transitive verb (e.g. *enjoys*) ($VP[trans:+]\{?*, \#1[last:+]\}$) can be recognised within a sentence chunk (SC) which is followed by an NP chunk not describing a time complement ($NP[time:\sim]\{?*, \#2[last:+]\}$). If this is the case, a VCOMP (with feature $dir:+$, for direct verb complement) dependency can be defined between the words at nodes $\#1$ and $\#2$, namely VCOMP[dir:+](wine,enjoy).

(Ait-Mokhtar et al., 2002) report the result of the evaluation of the linguistic performance of the XIP parser with the grammar of French. The system was tested with 7300 sentences taken from the newspaper *Le Monde*, for subject dependency and direct verb complements the precision obtained 93.45% and the recall 89.36%. (Trouilleux, 2001) has evaluated the same system for co-reference resolution obtaining a precision of 81.95% and a recall of 79.95%.

4.2 Incorporating VerbNet into XIP

To support a robust and large-scale treatment of alternation paraphrases, I extended XIP with VerbNet information and with a semantic construction module that assigns alternation paraphrases one and the same semantic representation. Briefly, the idea is to:

- integrate VerbNet information into a XIP lexicon, and
- to specify dependency rules which use this information (together with the VerbNet set of lexico-syntactic patterns) in order to assign a given input sequence the thematic grid assigned to that sequence by VerbNet.

In what follows, I present the lexicon and the semantic construction module I added to XIP.

4.2.1 The Verb Lexicon

To integrate VerbNet information into XIP, I specified a lexicon which associates each verb with its VerbNet class and with the WordNet Synset identifier corresponding to the relevant usage of that verb in that VerbNet class. Thus, a XIP verb lexicon entry has the following format:

```
verb_item :  verb += [VerbNet_Class=+, pred=WordNet_Synset]
```

where:

- `verb_item` is the particular verb entry,
- `verb` specifies the grammar category of the specific lexical entry,
- `+=`[...] is the feature list assignment for the lexical entry,
- `VerbNet_Class= +` is the name of the VerbNet class to which the verb entry belongs,
- `WordNet_Synset=` is the identifier for the WordNet synset to which the verb entry belongs.

Figure 4.5 shows XIP lexical entries for some verbs.
The VerbNet class is used both to guide syntactic parsing and to support semantic construction. Thus, as we shall see in the following section, only those dependency rules whose antecedent mentions the semantic class of the input verb will be triggered. Further, the VerbNet semantic class is used in the rule to specify the syntax/semantic interface, that is, the pairing between syntactic and semantic arguments.
The WordNet synset information, on the other hand, serves to group together synonyms. That is, all verbs in a VerbNet class which belong to the same WordNet synset will be assigned the same semantic representation. So, for instance, the verbs *meander, wander,*

```
abandon  :  verb += [leave51_2=+, pred=c02163637 ].
abash    :  verb += [amuse31_1=+, pred=c01740167 ].
...
break    :  verb += [break45_1=+, pred=c00323679 ].
break1   :  verb += [appear48_1_1=+, pred=c00543113 ].
break2   :  verb += [cheat10_6=+, pred=c02600809 ].
break3   :  verb += [split23_2=+, pred=c01260342 ].
break4   :  verb += [hurt40_8_311=+, pred=c00104829 ].
...
meander  :  verb += [ meander_47.7=+, pred = c01828635 ].
...
wander   :  verb += [meander47_7=+, pred=c01828635 ].
...
weave    :  verb += [meander47_7=+, pred=c01828635 ].
```

Figure 4.5: Xip Verb Lexicon

weave, wind and *thread* in the VerbNet class *meander-47.7* will all be assigned semantic information identical to that assigned to *meander*, i.e. the same WordNet synset (`pred=c01828635`).

The VerbNet class and synset assignment were made automatically on the basis of both VerbNet and WordNet information. At present, the lexicon contains 4229 verbs (all verbs contained in VerbNet) (corresponding to about 3010 disambiguated verbs, i.e. not ambiguous for class membership) corresponding to 2779 WordNet synsets and 352 VerbNet verb classes. However, since word sense disambiguation is not integrated in XIP, I only consider the most frequent meaning of a verb. This could be improved by integrating into XIP a verb sense disambiguation module such as the Word Sense Disambiguation System (Banerjee and Pedersen, 2003). Indeed, this module assigns each verb a WordNet synset number, which makes it directly compatible with my approach.

4.2.2 Extracting Argument Structures for English Verbs from VerbNet

VerbNet describes the argument structures possible for English verbs through 68 different syntactic frames. In order to facilitate the use of such information in XIP semantic rules, each of these syntactic frames was mapped to a more abstract representation, namely to the set of XIP grammatical functional dependencies corresponding to the frame.
The mapping between a VerbNet syntactic frame and its representation as set of XIP

HEAD	marks the head of a chunk	
Example	*The three girls*	HEAD(girls,The three girls)
	really nice	HEAD(nice,really nice)
	on the table	HEAD(table,on the table)
MAIN	marks the main verbal element of the principal clause of the parsed sentence.	
Example	*I wonder whether Mary will come to the party tonight.*	MAIN(wonder)
SUBJ	marks the subject of the sentence, it can have feature `pre` or `post` depending on whether it is on the left or on the right of the verb.	
Example:	*John is running in the park*	SUBJ[pre](running,john)
OBJ	marks the direct object of a verb	
Example	*I gave Mary some flowers*	OBJ(gave,flowers)
IOBJ	marks the indirect object of a verb	
Example	*I gave Mary some flowers*	IOBJ(gave,mary)
VMOD	marks a verb governor to any kind of complements or adjuncts attached to it, it can have feature `pre` or `post` depending on whether the modifier is on the left or on the right of the verb.	
Example:	*The new version will combine index with customised services*	VMOD[post](combine,services)
VDOMAIN	links the first and the last element of a complex verbal form, it can have feature `passive` or `modal`	
Example:	*Sara would like to sleep*	VDOMAIN[modal](like,would)
NMOD	marks a noun governor to any kind of complements or adjuncts attached to it, it can have feature `pre` or `post` depending on whether the modifier is on the left or on the right of the noun.	
Example:	*The inspector analysed the building's soundness.*	NMOD[pre](soundness,building)
NUCL_SUBJCOMP	marks the complement of a copula	
Example:	*Despite the warning, there was a unanimous vote to enter that candidate.*	NUCL_SUBJCOMP(was,vote)
SUBJATTR	marks the relation between the subject and the subject complement of a copula	
Example:	*Despite the warning, there was a unanimous vote to enter that candidate.*	SUBJATTR(there,vote)
PREPD	link between a preposition and the nominal head of a PP, it can have feature `prep` to constrain the preposition	
Example	*I gave flowers to Mary*	PREPD[prep:to](to,mary)

Figure 4.6: Some XIP Dependencies describing grammatical functions

dependency relations was done automatically. For each syntactic frame the example sentences in VerbNet were parsed with XIP and the most frequent dependency representation output by XIP was taken to be the correct one and associated to the correspondent VerbNet frame as its abstract representation.
A description[14] of the grammatical functions implemented in the English grammar provided by XIP is given in Figure 4.6.
The semantic rules will then apply to such representations of the syntactic frames. Some of the mappings defined between VerbNet frames and XIP functional dependencies are described in Figure 4.7.

4.2.3 Semantic Construction

To assign identical semantic representations to alternation paraphrases, I have extended the XIP grammar with a set of *thematic grid dependency rules*[15]. These rules assume as input the output of the existing XIP parser, that is, a representation of the input including both constituency and grammatical functions (subject, object, etc.) information. Based on this information, a thematic grid (dependency) rule identifies a given VerbNet pattern (syntactic frame and verb semantic class) and specifies a mapping between syntactic and thematic arguments. In fact, the same argument structure can be mapped to a different grid of theta roles depending on the VerbNet class membership of the verb under consideration.
Figure 4.8, for example, shows two possible VerbNet mappings of a `Transitive + PP` argument structure. If the verb which accepts this frames belongs to the *give-13.5.1* class, the subject NP will be mapped to the `AGENT` role, the object NP to the `THEME` role and the PP to the `BENEFICIARY` role. On the other hand, if the verb belongs to the *get-13.1.1* class, the subject NP will be mapped to the `AGENT` role, the object NP to the `THEME` role but the PP to the `RECIPIENT` role.
Let me now illustrate with a simple example how semantic construction is performed. Suppose the sentence to be parsed is:

(60) *The river meanders through the valley*

As we know from the last chapter, the VerbNet syntactic and semantic information associated with that usage of the verb *meander* is:

VerbNet class	*meander47-7*
Syntax:	`Theme V Prep[+path] Location`

[14]This description is taken from XIP English Grammar User's Guide of XIP version 3.10

[15]These rules can be seen as rewriting rules as the metarules (Uszkoreit and Peters, 1986) of GPSG. However, while the GPSG metarules capture long distance dependencies and handle with missing constituents, the semantic rules described here rewrite grammatical functions (which should be present) to theta roles.

VerbNet Frame	XIP Representation
Transitive + PP	
[NP$_1$, V, NP$_2$, PP]	MAIN(V), SUBJ(V,$\mathcal{H}ead$(NP$_1$)), OBJ[post](V,$\mathcal{H}ead$(NP$_2$)), VMOD(V,$\mathcal{H}ead$(PP))
Dative Ditransitive Variant	
[NP$_1$, V, NP$_2$, NP$_3$]	MAIN(V), SUBJ(V,$\mathcal{H}ead$(NP$_1$)), OBJ[post](V,$\mathcal{H}ead$(NP$_2$)), OBJ[post](V,$\mathcal{H}ead$(NP$_3$))
Intransitive with PP	
[NP, V, PP]	MAIN(V), SUBJ(V,$\mathcal{H}ead$(NP)), ~OBJ(V,?), VMOD[post](V,$\mathcal{H}ead$(PP))
Intransitive with PP1 and PP2	
[NP, V, PP$_1$, PP$_2$]	MAIN(V), SUBJ(V,$\mathcal{H}ead$(NP)), ~OBJ(V,?), VMOD[post](V,$\mathcal{H}ead$(PP$_1$)), VMOD[post](V, $\mathcal{H}ead$(PP$_2$))
Middle Construction	
[NP, V, Adv]	MAIN(V), SUBJ[pre](V,$\mathcal{H}ead$(NP)), VMOD[post](V,Adv), ~PREPD(?,Adj)
Resultative Construction	
[NP, V, Adj]	MAIN(V), SUBJ[pre](V,$\mathcal{H}ead$(NP)), NUCL_SUBJCOMPL(V,Adj), SUBJATTR($\mathcal{H}ead$(NP), Adj)
NP-ADJP Resultative Construction	
[NP$_1$, V, NP$_2$, Adj]	MAIN(V), SUBJ[pre](V,$\mathcal{H}ead$(NP$_1$)), OBJ(V,$\mathcal{H}ead$(NP$_2$)), NMOD[post]($\mathcal{H}ead$(NP$_2$), Adj)

Figure 4.7: Mapping between VerbNet Syntactic Frames and XIP Dependency Representation

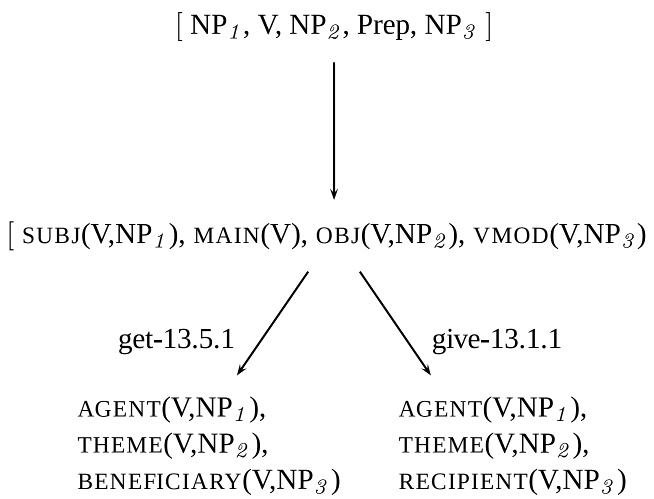


Figure 4.8: Verbnet Mappings of a `Transitive+PP` argument structure to theta roles

The syntactic description abbreviates the following specification: the canonical subject is a theme and a prepositional object introduced by a path preposition which denotes the location of the event. In the XIP framework, such a specification is captured by the (simplified) dependency rule shown in Figure 4.9. This rule is a standard XIP dependency rule with an empty subtree pattern, a condition and a list of dependency terms as output. In this rule, || denotes disjunction, MAIN, VDOMAIN, SUBJ, PREDP and VMOD are dependencies as explained in Figure 4.6.

In other words: if the main verb (#1) is associated (via lexical lookup) with one of the listed VerbNet classes (e.g. *coil_9_6* or *coil_9_61* or *escape_51_1* or *escape_51_11*, etc. and in particular with the *meander47_7* class), if it is not in the passive mode (`VDOMAIN[passive:~](#1,#11)`) and has no object (`~OBJ(#1,?)`) but has a subject (`SUBJ(#1,#2)`) and a postposed verb modifier (`VMOD[post](#1,#4)`) introduced by a path denoting preposition (`PREDP(#3[vnpath],#4)`), then the semantic representation produced is

```
EVENT(#1), THEME(#1,#2),LOCATION(#1,#4)
```

where #1, #2 and #4 are the nodes associated with the main verb, the subject and the modifier head respectively.

As this rule applies to the input sequence (60), the following representation is output where indeed the correct thematic representation has been assigned to the sentence.

```
if( (  MAIN(#1[coil9_6])
    || MAIN(#1[coil9_61])
    || MAIN(#1[escape51_1])
    || MAIN(#1[escape51_11])
    || MAIN(#1[escape51_12])
    || MAIN(#1[escape51_121])
    || MAIN(#1[meander47_7])
    || MAIN(#1[substance_emission43_4])
    || MAIN(#1[vehicle51_4_1])
    || MAIN(#1[vehicle51_4_11])
    || MAIN(#1[waltz51_5])
    )
     & VDOMAIN[passive:~](#1,#11)
     & SUBJ(#1,#2) & ~OBJ(#1,?)
     & VMOD[post](#1,#4) & PREPD(#3[vnpath],#4)
  )
EVENT(#1),THEME(#1,#2),LOCATION(#1,#4).
```

Figure 4.9: A semantic construction rule

(61) `EVENT(meanders), LOCATION(meanders,valley),`
`THEME(meanders,river)`

More generally, the extended XIP grammar counts 425 semantic rules. These rules are ordered by specificity, the most specific rules occurring first and the least specific last. For instance, the rules for ditransitives will be tested before the rules for transitives which again will be tested before the rules for intransitives. Since within one grammar layer only the first applicable rule is used, this ensures that the syntactic configuration captured by the rule that is executed is indeed the most appropriate. This rule ordering also allows an appropriate treatment of the difference between adjuncts and subcategorised PPs. Being more specific, the rules describing verbs taking prepositional arguments will be tested before the general rules describing the combination of verbs with adjuncts and so will be preferred in case they can be applied. Here is an illustrating example. Suppose we have the two sentences given in (62) and we want to obtain the indicated semantic representations.

(62) a. Sharon shivered from fear.
`EVENT(shivered), CAUSE(shivered,fear),`
`EXPERIENCER(shivered, Sharon)`

b. Sharon breakfasted in the garden.
`EVENT(breakfasted), AGENT(breakfasted,Sharon)`

In the first sentence (62a), the PP is described in VerbNet as an element of the subcategorisation frame of the verb *shiver*, which is mapped to the `CAUSE` role. In contrast, in

```
**** RULE 106: ****
if( ( MAIN(#1[body_internal_states40_6])
    || MAIN(#1[change_bodily_state40_8_4])
    )
     & VDOMAIN[passive:~](#1,#11)
     & SUBJ(#1,#2) & ~OBJ(#1,?)
     & VMOD[post](#1,#4) & PREPD(#3[prep:from],#4)
  )
EVENT(#1),EXPERIENCER(#1,#2),CAUSE(#1,#4).
...
*** RULE 419: ***
if( ( ... || MAIN(#1[dine39_5]) || ... )
     & VDOMAIN[passive:~](#1,#11)
     & SUBJ(#1,#2) & ~OBJ(#1,?)
  )
EVENT(#1),AGENT(#1,#2).
```

Figure 4.10: Semantic Construction Rules for *shiver* and *breakfast*

the second sentence (62b), the PP is treated as an adjunct and is not assigned a thematic role. By placing the rule describing the “shiver” configuration (Rule 106) before the more general rule describing intransitive patterns, we can ensure that both sentences are assigned the correct thematic grid. In case the input is (62a), the “shiver” rule is first to apply, thereby licensing the construction of the given semantic representation. For (62b) on the other hand, the “shiver” rule does not apply (because the verb *breakfast* does not belong to the same VerbNet class as *shiver*) but the general intransitive rule does, which fails to assign the PP a thematic role[16].

To define the specificity ordering over the thematic rules, I first generalised the syntactic frames to 68 more general templates by ignoring prepositional and selectional information. For instance, the VerbNet syntactic frames in (63) were both abstracted to the more general template `NP V PP`.

(63) *`NP V Prep(of) NP`*
`NP V Prep(with) NP`

The resulting set of templates was then organised in a hierarchy (cf. Figure 4.11[17]) which was then used to automatically order the XIP thematic rules.

[16]Of course the locative PP *the garden* should be assigned a semantic representation too and be related, e.g. by a locative relation to the described event. I do not discuss this here as I am only concerned with correctly describing the thematic roles of the *arguments* of a verb.

[17]The depicted hierarchy presents a fragment of the whole hierarchy which, in total, describes 68 verb/arguments patterns.

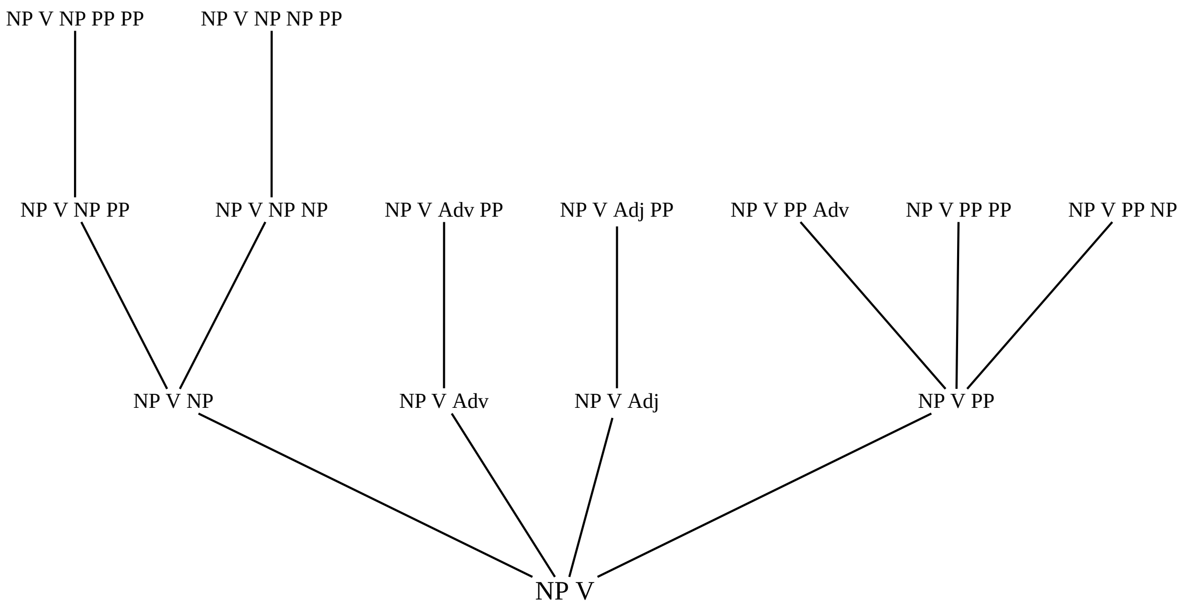


Figure 4.11: Hierarchy of syntactic patterns in XIP

4.2.4 Postprocessing

To cover unknown input and more specifically verbs whose VerbNet class is not given in the lexicon, I introduce an additional postprocessing step which performs a default thematic grid assignment on the basis of the 68 syntactic frames extracted from VerbNet as described in the previous section Specifically, these very general rule templates are used to specify 68 general rules describing general subcategorisation frames and assigning a default role to each of the arguments identified through those frames. For instance, suppose that the input sentence is (64) and that the VerbNet class for *stand* is not given in the lexicon.

(64) *On the pedestal stood a statue*

In such a case, the general rule specifying a syntactic configuration of the form `[PP[+loc] V NP]`, will assign the locative PP an `arg2` role and the NP an `arg1` role, thereby producing the semantic representation given below.

```
EVENT(stood), ARG1(stood,statue),
ARG2(stood,pedestal)
```

4.3 A More Detailed Example

Consider the paraphrastic sentences in (65). If such sentences are parsed by the original version of XIP, the result is the two different (deep-)syntactic analyses given in Figure 4.12.

(65) a. *I rented my apartment to Ann.*

b. *I leased Ann my apartment.*

INPUT
S1: *I rented my apartment to Ann.* **S2:** *I leased Ann my apartment.*
NTM, HHM
S1: I_`Pron` rented_`Verb` my_`Pron` apartment_`Noun` to_`Prep` Ann_`Noun` **S2:** I_`Pron` leased_`Verb` Ann_`Noun` my_`Pron` apartment_`Noun`
Chunking Rules
S1: `TOP [SC [NP I] [FV rented]] [NP my apartment] [PP to Ann]` **S2:** `TOP [SC [NP I] [FV leased]] [NP Ann] [NP my apartment]`
Dependency Rules OUTPUT
S1: MAIN(rented) SUBJ(rented,I) VMOD(rented,Ann) OBJ(rented,apartment) **S2:** MAIN(leased) SUBJ(leased,I) OBJ(leased,Ann) OBJ(leased,apartment)

Figure 4.12: XIP syntactic analysis

Both verbs *sell* and *lease* belong to the VerbNet class *give-13.1.1* described in Figure 4.13 and are synonyms. The argument structures of these verbs and their mapping to theta roles are given below[18]:

- Transitive + PP: [NP$_1$, V, NP$_2$, Prep, NP$_3$]
  XIP Dependency Representation:
  MAIN(V), SUBJ(V,NP$_1$), OBJ(V,NP$_2$), VMOD(V,NP$_3$)
  VerbNet Mapping to Theta Roles:
  `AGENT(V,NP`$_1$`), THEME(V,NP`$_2$`), RECIPIENT(V,NP`$_3$`)`

- Dative Ditransitive Variant: [NP$_1$, V, NP$_2$, NP$_3$]
  XIP Dependency Representation:
  MAIN(V), SUBJ(V,NP$_1$), OBJ(V,NP$_2$), OBJ(V,NP$_3$)
  VerbNet Mapping to Theta Roles:
  `AGENT(V,NP`$_1$`), THEME(V,NP`$_3$`), RECIPIENT(V,NP`$_2$`)`

This information is captured in the XIP semantic construction rules depicted in Figure 4.14. By applying these rules, the XIP extended parser assigns to the two different syntactic frames possible for the two verb classes the same theta roles and thus to the sentences in (65) the same thematic grid.

[18]For better readability, I have simplified the XIP representations e.g., by omitting the node number and directly stating dependencies between words rather than between nodes.

The VerbNet class GIVE-13.1.1
Members: *give(1), rent(1), sell(1), lease(1)*
Thematic Roles and Selectional restrictions: Agent[+animate OR +organisation] Theme[] Recipient[+animate OR +organisation] Asset
Frames: `Transitive (+ Recipient PP)` "They lent a bicycle to me" `Agent V Theme Prep(to) Recipient` *has_pos(start(E),Agent,Theme) transfer(during(E),Theme) cause(Agent,E) has_pos(end(E),Recipient,Theme)*
`Dative (Ditransitive Variant)` "They lent me a bicycle" `Agent V Recipient Theme` *has_pos(start(E),Agent,Theme) transfer(during(E),Theme) cause(Agent,E) has_pos(end(E),Recipient,Theme)*

Figure 4.13: VerbNet representation of the *give-13.1.1* class

```
**** RULE 189: Ditransitive Variant ****
| NP{?*,\#3[last]}, NP{?*, \#4[last]} |
if( ( ... || MAIN(\#1[give-13.11]) || ... )
     & VDOMAIN[passive:~](#1,#11)
     & SUBJ(#1,#2) & OBJ[post](#1,#3)
     & OBJ[post](#1,#4)
  )
EVENT(#1), AGENT(#1,#2), THEME(#1,\#4), RECIPIENT(\#1,\#3).
...
*** RULE 294: Transitive + Recipient PP ***
if( ( ... || MAIN(\#1[give-13.11]) || ... )
     & VDOMAIN[passive:~](#1,#11)
     & SUBJ(#1,#2) & OBJ[post](#1,#3) &
     & VMOD[post](#1,#5) & PREPD(#4[prep:to],#5)
  )
EVENT(#1), AGENT(#1,#2), THEME(#1,\#3), RECIPIENT(\#1,\#5).
```

Figure 4.14: Semantic construction rules for *give-13.1.1* class

S1: `EVENT(rented) AGENT(rented,I) RECIPIENT(rented,Ann) THEME(rented, apartment)`
S2: `EVENT(leased) AGENT(leased,I) RECIPIENT(leased,Ann) THEME(leased, apartment)`

In the semantic construction step, the output of the XIP dependency grammar analysis of the sentences in (65) (see Figure (4.12)) is matched against semantic construction rules.
Rule 189 only applies to sentence (65a), as this rule constrains two adjacent NPs ($|NP\{?*, \#3[last]\}, NP\{?*, \#4[last]\}|$) following the main verb #1 to be both its objects ($OBJ[post](\#1, \#3), OBJ[post](\#1, \#4)$). Rule 294 only applies to sentence (65b), as it constrains the main verb #1 to have an object ($OBJ[post](\#1, \#3)$) and to be further modified by a PP ($VMOD[post](\#1, \#5)$) with preposition *to* ($PREPD(\#4[prep : to], \#5)$).

4.4 Evaluation

The evaluation aimed to assess the degree to which the extended XIP parser could capture the kind of verbal alternations described in Levin's work and more specifically, in Verbnet. The aim was to evaluate whether the system did indeed produce the thematic grid expected by VerbNet for a given verb and a given syntactic configuration.

To carry out this evaluation, I extracted for each VerbNet class the example sentences it contains together with their thematic role annotation. For instance, (66) shows the annotated sentences included in the *meander-47.7* class.

(66) a. *The river runs through the valley*
`THEME V Prep[+path] LOCATION`

b. *Through the valley meanders the river*
`Prep[+path] LOCATION V THEME`

c. *There meanders through the valley a river*
`there V Prep[+path] LOCATION THEME`

d. *There meanders a river through the valley*
`there V THEME Prep[+path] LOCATION`

Out of the 361 classes contained in VerbNet[19], I extracted 1012 example sentences illustrating the various frames of each of these 361 classes.

I then applied the parser to this set of sentences and automatically compared the thematic grid output by the extended XIP parser with the thematic grid described by the VerbNet annotation. XIP assigns to 90% of the 1012 example sentences extracted

[19]The numbers given here concerning VerbNet data back to the time the work presented in this chapter was carried out namely, 2005. Since then the numbers have changed so that there are now for instance, not 361 but 471 classes in VerbNet.

from VerbNet the expected syntactic representation, i.e. the one matching the expected Verbnet syntactic frame.

Ignoring the sentences which yield an incorrect syntactic parse, I obtained the following results by considering 911 sentences and by using a Xip lexicon which includes 3010 verbs:

- 79% of the sentences were assigned the correct representation (i.e. the same roles assignment as in VerbNet),
- 17% of the sentences were assigned the correct syntactic pattern but the wrong theta roles because selectional restrictions could not be checked
- 4% of the sentences were assigned a default pattern because either the tagger was not able to recognise the class membership of the verb, or the verb class assignment in the lexicon did not allow the syntactic pattern illustrated by the given sentence.

The problem with selectional restrictions can be illustrated by the following example. In VerbNet the verb *buy* is assigned (among others) the following two frames:

(67) *a.* *Basic Transitive*
Carmen bought a dress
`AGENT V THEME`

b. *Sum of Money Subject Alternation (Asset Subject)*
$50 won't even buy a dress at Bloomingdale's
`ASSET V THEME`

Without knowledge about the ontological type of the arguments, it is impossible for the parser to decide whether the subject of the sentence should be assigned an `AGENT` or an `ASSET` thematic role. In other words, for 17% of the VerbNet data ontological knowledge is required in order to correctly determine the thematic grid of the input sentence. Such knowledge could be integrated into XIP by resorting e.g. to the WordNet hierarchy which is linked to the verb usages described by VerbNet.

The remaining 4% requires improving both the tagger (in case the verb is not tagged appropriately) and the VerbNet description (in case a syntactic pattern is missing for a given verb usage). Further, as the lexicon includes only the most frequent reading of a verb, if different readings of a verb belong to two different classes and have different syntactic frames, one of them will not be captured. To solve this shortcoming a word disambiguation module is required.

In sum, the evaluation shows that the parser developed deals appropriately with 79% of the VerbNet data and that it could assign correct results to 86% of the example sentences extracted from VerbNet, if it is extended with ontological knowledge. Thus, there is reason to hope that it can be further improved by incorporating selectional restrictions, a word sense disambiguation and by improving the basic constituency grammar.

Moreover, the postprocessing step allows the parser to assign default underspecified thematic roles to maximal projection phrases occurring in previously unseen input on the basis of surface syntax information. In such cases, the use of such underspecified thematic roles renders the obtained semantic representations similar to those assumed by PropBank (Kingsbury et al., 2002).
Nonetheless, the evaluation benchmark remains limited. Benchmarks such as the PropBank or the ConLL data on semantic role labelling provide real life, textual data in which verb/arguments dependencies have been manually annotated. For instance, the test corpus of PropBank provides a set of 8248 annotated sentences on which to evaluate and compare semantic role labelling systems. It would be interesting to evaluate and improve the extended XIP parser on such data.

4.5 Summary

In this chapter, I presented a linguistic-based computational approach to verb paraphrasing. I have shown that the detailed knowledge about verb alternations encoded in VerbNet can automatically be integrated in a robust parser like XIP, thereby supporting the recognition of the set of alternation and/or lexically synonymic paraphrases covered by VerbNet.
In the next chapter I compare the most relevant related works in this field, linguistic- and statistic-based approaches, with the approach presented in this chapter.

Related Work

Recent years have shown a strong interest in paraphrases, as they represent a crucial issue in many NLP applications. In Question Answering applications, for instance, it is crucial for a system to have the ability to recognise different verbalisations of a question in order to individuate documents which contain the answer. In Information Retrieval, paraphrasing is used to perform query expansion, i.e. a query is expanded with all its possible variations. In multi-document summarisation, paraphrase recognition is important in order to decide which text-snippets sharing the same content should be used in the summary.

Much of the work recently done on paraphrases has focused especially on information extraction (IE) and question answering (QA). And because IE and QA systems deal with large, open domain corpora, much of the work on paraphrases is based on automatic learning techniques. However, there are also some interesting linguistic-based approaches to paraphrasing.

In the first part of this thesis, I proposed a symbolic, linguistic-based treatment of verb paraphrases which is robust and has wide coverage. In this chapter, I compare my approach with the statistical approaches used in information extraction and question answering as well as with some related symbolic approaches.

5.1 Statistical Approaches

In the following, I give an overview of statistical approaches. Such approaches generally propose a strategy to automatically learn paraphrastic patterns from unseen texts.

Similarity Measures on Trees

Lin and Pantel (2001) propose an unsupervised algorithm called DIRT to discover inference rules (i.e. paraphrastic patterns) from texts. The DIRT algorithm applies an extended version of (Harris, 1985) Distributional Hypothesis to paths of dependency trees obtained after parsing a text. Paths in dependency trees that have similar arguments are taken to be close in meaning and to represent binary semantic relations. Thus, similar paths are used to build inference rules. Consider, as an example, the dependency trees a) and b) in Figure 5.1 which correspond to the analysis of the paraphrastic sentences below:

(68) a. *John found a solution to the problem.*

b. *John solved the problem.*

The trees a) and b) have in this case identical arguments, thus they are used to generate the paraphrastic pattern or inference rule in (69).

(69) $\texttt{X find solution\_to Y} \approx \texttt{X solve Y}$

This algorithm has been evaluated by comparing the dependency tree representations of a set of questions, obtained by generating paraphrases for the first 6 questions of the (TREC-8, 1999) Question Answering Track, with those of the paraphrastic paths extracted from a corpus of newspapers.
The results of the evaluation show that while in some cases the DIRT algorithm can find more paraphrastic structures or relations than human annotators, it still displays limits typical of statistically-based approaches. Namely the coverage is limited: (i) if the path contained in the question was not in the database then no paraphrase is found, and (ii) the algorithm only copes with binary relations.

A more fundamental limit of this approach is that it cannot cope with polarity. So for instance, the paths in Figure 5.1 are judged to be paraphrastic although *worsen* and *solve* have meanings which in some way refer to opposite situations, if one considers that after solving a problem the situation is better than before.

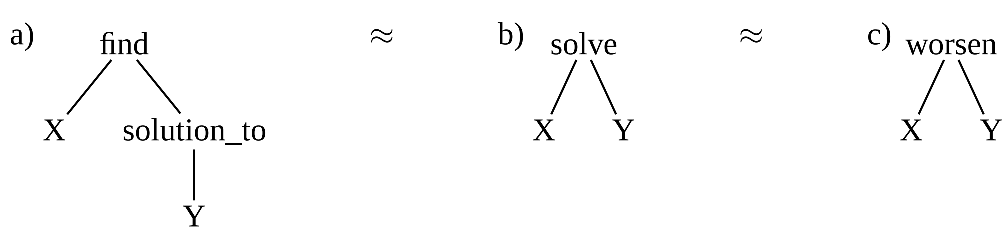


Figure 5.1: DIRT Inference Rules

Automatic Paraphrase Extraction

Barzilay and Lee (2003) and Shinyama et al. (2002) learn sentence level paraphrase templates from a corpus of news articles on the same topic but stemming from different

news sources.
Barzilay and Lee (2003) acquire recurrent patterns as follows. First, structurally and semantically similar sentences stemming from comparable corpora are clustered by applying hierarchical complete-link clustering to sentences using a similarity metric based on word n-gram overlap (n=1;2;3;4). To minimise the impact of minor, surface level differences, arguments with great variability, such as locations, names and number expressions are substituted with slots. In a second step, phrasal paraphrase patterns such as L1 and L2 in (70) are learnt from these clusters using multiple sequence alignment (MSA) of the sentences in a given cluster.

(70) L1 : slot1 bombed slot2

L2 : slot3 was bombed by slot4

The paraphrase patterns are then used to generate paraphrases by finding, for a given sentence, the most similar phrasal patterns found by the system and substituting in the other member of the paraphrase pattern the appropriate input sentence phrases.

On a set of 100 paraphrase patterns extracted randomly from the system output, four judges evaluated the validity of the extracted patterns and estimated it to be varying between 68% , 74% , 76% and 96% respectively.

Starting from the idea that Named Entity (names, dates, numbers, etc.) are constants in paraphrases, Shinyama et al. (2002) acquire paraphrastic template from newspaper articles describing the same event. Similarity measure on sentences is done on the basis of the number of the shared NEs. The larger the number of shared NEs, the larger is the probability that the two sentences are peraphrastic.

The result of the evaluation shows that while in some domain (e.g. (71)) the algorithm performs well in others (e.g. (72)), it does not. In fact, the frequency of NEs depends of the particular domain.

(71) P1 : ORGANISATION decides

P2 : ORGANISATION confirms

(72) P3 : is promoted to POST

P4 : POST is promoted

Glickman and Dagan (2003) use clustering and similarity measures to identify similar contexts in a single corpus and extract verbal paraphrases from these contexts. In this approach, during parsing, the text is divided into sentences which contain only one verb. Sentences are represented in a vector which contains as a main component the verb head of the sentence and information about its arguments. *Verb instance pairs*, i.e. pairs of vector sentences, are then compared for their argument representation, if two verb instances share the same subject and object they are considered to be likely paraphrases.

Verbal paraphrases are then extracted from the corpus by using overlap/similarity measures and paraphrase likelihood computed for each verb instance pair.

One shortcoming common to approaches which use syntactic overlapping measures to judge whether two structures are paraphrastic or not stems from antonymic and converse constructions. This is because, as illustrated in (73) and (74), antonymy and converse constructions tend to occur in similar syntactic structures.

(73) a. Rand Financials notably bought October late while locals lifted December into by stops

b. Rand Financials notably sold October late while locals pressured December

(74) a. French shares opened lower due to renewed pressure on the franc ...

b. French shares closed lower due to a weaker franc ...

On this particular data, Glickman and Dagan (2003) for instance, identifies the pairs of verbs $\langle open, close\rangle$ or $\langle buy, sell\rangle$ as paraphrases or synonyms although the verbs of the first pair are clearly antonyms and those in the second are converses.

Paraphrase for Restricted Sets of Question-patterns

In the context of open domain question answering systems, Ravichandran and Hovy (2002) use bootstrapping techniques to learn paraphrastic text patterns. The algorithm makes use of question types. For each question type, a question-answer pair is used to search the web for patterns: from question and answer, a pair of terms is extracted containing the answer and the question, for instance from the question-answer pair ⟨'When was Mozart born?'/1756⟩ the pair of terms *Mozart* and *1756* is extracted. These two terms are used as queries to search the web for documents containing the two strings. From these documents, phrases containing the terms are extracted and rewritten as regular expressions. By substituting the terms in regular expressions with variables, a pattern is generated. The precision scores of the obtained patterns are calculated by comparing these patterns with the patterns found in the web by using as query only the question term. An example of the paraphrastic patterns extracted with this method for the question type BIRTHDATE is given in Figure 5.2. This method while allowing a quick learning of numerous paraphrastic patterns cannot capture more complex phenomena such as for example, long-distance dependencies or definitions. Further it is prone to overgenerate that is, to generate illicit paraphrastic patterns, as the regular expressions used to describe the patterns representing a question-answer pair are often necessarily very general.

Question Type: BIRTHDATE
a. born in $\langle ANSWER\rangle$, $\langle NAME\rangle$ b. $\langle NAME\rangle$ was born on $\langle ANSWER\rangle$ c. $\langle NAME\rangle$ ($\langle ANSWER\rangle$ - d. $\langle NAME\rangle$ ($\langle ANSWER\rangle$ -)

Figure 5.2: Paraphrastic patterns extracted for question type BIRTHDATE

5.2 Symbolic Approaches

Combining VerbNet with XTAG

Kipper et al. (2000); Ryant and Kipper (2004) propose an integration of VerbNet with the XTAG (see Group, 2001) grammar, a lexicalised Tree Adjoining Grammar of English which covers the possible transformations of canonical frames. Specifically, a mapping is specified between VerbNet frames and XTAG tree families whereby each VerbNet frame is mapped to a corresponding XTAG tree family. Since an XTAG tree family specifies the possible transformations (active, passive, extrapositions, etc.) of a given syntactic frame, this mapping in effect extends the coverage of VerbNet beyond canonical frames to all transformations of these canonical frames.

In itself, the mapping provided by Kipper et al. (2000); Ryant and Kipper (2004) does not suffice to support alternation based paraphrases. A processor additionally needs to be specified which uses this mapping to automatically convert the XTAG parses to the kind of more abstract, semantic representations needed to establish a paraphrastic link between two sentences. It could be used however, to extend the coverage of the XIP approach presented here to non canonical variants. Indeed, the approach was only evaluated on canonical sentences that is, on sentences in the active form and with default argument order. Whilst it is possible that the treatment of syntactic variations provided by XIP results in ascribing non canonical variants a dependency structure compatible with the semantic construction rules determining the mapping between syntactic and semantic arguments, it is rather unlikely. To remedy this shortcoming, the mapping defined by Kipper et al. (2000); Ryant and Kipper (2004) could be used to support the automatic derivation of semantic construction rules for non canonical variants from semantic construction rules for canonical ones.

An Alternative XIP-based Approach

Brun and Hagège (2003); Hagège and Roux (2003) also describe an attempt to extend XIP to recognise paraphrases. They use hand-written rules in order to map the normalised output of XIP to a more fine grained semantic representation, capturing alternations, synonymy and derivational morphology.

They explain in detail how they capture synonymy and nominalisations, i.e. the pair of paraphrastic constructions in which a verb is substituted with a morphologically related noun. For example,

(75) a. *Mary plays the violin*

b. *Mary plays the fiddle*

(76) a. *Smith invented the process*

b. *Smith is the inventor of the process*

In order to handle such cases of paraphrases, information about synonymy is introduced as a set of dependency relations (Syn), each of which maps a word to one of its synonyms. For example:

```
Syn(violin,fiddle)
```

Derivational morphology information is introduced in the form of the four relations between noun and verb described in Figure 5.3. Then, hand written rules test whether such (synonymy and/or derivational) relations hold between some words of the current input. The rule below, for example, can extract the OBJ and SUBJ dependencies from syntactic structures such as the one in sentence (76b) by making use of derivational morphology dependencies:

Lexicon
S1H(inventor,invent)
Extraction Rule
if (S1H(#1,#9) & MOD_PREP[form:of](#1,#2) & ATTRIB(#1,#3))
OBJ(#9,#2), SUBJ(#9,#3)
Extraction
OBJ(invent,process), SUBJ(invent,Smith)

The problem of this approach is that information which should be stored in the lexicon is coded as dependency, making the formalism nonmodular and often inefficient. In the parser developed in this thesis, synonymy is captured directly during analyses, as the meanings of verbs are passed as features to semantic relations, so that no further test is needed.
Moreover, Brun and Hagège (2003); Hagège and Roux (2003) do not provide an explanation of how their system captures alternations and of what kind of alternations they can recognise and no evaluation of the system is given.

S_0	the morphologically related noun expresses the same action or event of the verb	(ex. *acceleration/accelerate*)
S_{1H}	the related noun expresses the first semantic (agent) actant of the verb with feature human	(ex. *beginner/begin*)
S_{1NH}	the related noun expresses the first semantic (theme) actant of the verb without feature human	(ex. *abbreviation/abbreviate*)
S_2	the related noun expresses the second semantic actant of the verb	(ex. *affirmation/affirm*)

Figure 5.3: Nominalisations: dependency relations between noun and verb

Automatic Generation of Syntactically Well-Formed and Semantically Appropriate Paraphrases

Fujita (2005) proposes a lexical semantic-based approach to paraphrase generation focusing on paraphrasing of Japanese light verb constructions. He uses the framework of Lexical Conceptual Structure (Jackendoff, 1990) to represent verb semantics in a decompositional manner and to map verb arguments to their syntactic realisation. Relying on a set of hand coded transformation rules and on the LCS lexicon for Japanese verbs (Takeuchi, 2004), the generation system generates all possible candidate paraphrases for a given input sentence which in a later step are detected for errors by using linguistic resources and then corrected if possible or rejected.
The author says nothing about the coverage of his system. He uses hand-coded transformation rules to cope with paraphrases. On the contrary, the parser presented in this thesis makes sole use of linguistic information to build similar semantic representations for paraphrases, in this way being more robust.

Semantic Role Labelling

Recently, (Kaisser and Webber, 2007) have used VerbNet, FrameNet and PropBank in a Question Answering system setting to obtain from a question, a set of potential answer templates.

The first of the two methods they describe has many similarities with the work presented here and can be sketched as follows. First, the question is parsed and both its head verb V and its syntactic arguments are extracted. Next the information (frames, mapping syntax/semantics, annotated examples) contained in VerbNet, FrameNet and PropBank is used to construct a set of frames describing the possible syntax/semantic argument mapping for V. Next, the frame that fits the syntactic analysis best is selected and the

Input Sentence
Ken received an inspiration from the film

LCS_input
[BECOME$[[Ken]_z$ BE WITH $[[Inspiration]_y$ MOVE FROM$[film]_x$ TO $[Ken]_z]]]$

Transfer Rule
[BECOME $[z$ BE WITH $[y$ MOVE FROM x TO $z]]] \rightarrow [x$ ACT ON $y]$

LCS_output
$[[film]_x$ ACT ON $[Ken]_y]$

Output
Ken was inspired by the film

Figure 5.4: Paraphrasing of Japanese light-verb constructions

corresponding syntax/semantics mapping is applied to determine the semantic/thematic roles of V's syntactic arguments.

The approach is similar to the extended XIP approach in that is uses VerbNet (among others) to determine the thematic role of a verb constituent. As in my approach, the mapping between syntactic and semantic arguments is done by relating a syntactic configuration given by a parser to a semantic frame associated with the verb under consideration by a lexical resource. The system was evaluated in a Question/Answer system and was shown to substantially improve accuracy. Moreover Kaisser and Webber (2007)'s approach substantially improves on the XIP proposal by interfacing the dependency structures produced by parsing not only with VerbNet but also with Propbank and with Framenet. As a result, it achieves better coverage, can make use of the different levels of knowledge encoded in each resources (e.g., coarse thematic roles in PropBank, finer grained thematic ones in VerbNet and more semantically oriented roles in FrameNet) and can potentially handle not only verbal but also nominal alternations (because they are covered in Framenet).

No details are given in the paper on how the diverging representation schemes (in particular, the distinct set of thematic/semantic roles) used by the three resources are handled and/or merged. It would be interesting to explore this fusion in more detail and to investigate whether a common scheme can be specified which permits a generalised, symbolic treatment of alternation-based paraphrases. An intrinsic evaluation on the dataset I used would also be interesting that would permit comparing their approach with the extended XIP system.

5.3 Discussion

As we have seen in this chapter, statistical approaches to paraphrasing are mainly based on machine learning techniques that have known pros and contra. On the one hand, such methods produce large scale resources at little man labour cost. On the other hand, the degree of descriptive abstraction offered by the list of inference or paraphrase rules they output is very low. It is worth noting, for instance, that none of the statistical approaches described in this chapter can cope with the problem of polarity of lexical items, i.e. can recognise converse constructions or antonyms as having different meaning with respect to the positive form.

In summary, the difference between the statistical approaches presented above and the approach presented in this thesis are those generally occurring between symbolic and statistical approaches. First, the coverage differs. Whereas the approach presented here concentrates on alternation paraphrases and verbal synonymy, statistical approaches either take a very general way of identifying paraphrases grouping together e.g. synonyms, hyponyms and siblings (Barzilay and McKeown, 2001) or are specialised to a few chosen relations occurring frequently in a given domain (Ravichandran and Hovy, 2002). Second, the resources used differ in that statistical treatments of paraphrases rely on aligned related corpora, on corpora treating of the same topic or on extremely large corpora (the web). By contrast, the approach presented in this thesis is based on existing symbolic resources, namely VerbNet and WordNet. Third, the output of the two approaches differs in that statistical approaches typically yield a *paraphrase lexicon*, that is, a list of paraphrases which is independently put to work in a given application by some string manipulation procedure. In contrast, the output of the framework proposed here is a parser designed to handle alternation paraphrases.

On the other hand, symbolic approaches to paraphrasing propose to integrate detailed linguistic information in computational systems so to enhance their ability to cope with paraphrases. So for instance, Kipper et al. (2000); Ryant and Kipper (2004) try to build a compositional semantic framework able to recognise paraphrases by linking VerbNet verb classes to tree adjoining grammars (XTAG), while Brun and Hagège (2003); Hagège and Roux (2003), similarly to me, use XIP to assign to two paraphrases the same representation and Fujita (2005) proposes a lexical semantic-based approach to paraphrases generation based on Lexical Conceptual Structure. The use of tree adjoining grammar extends the coverage of the (Kipper et al., 2000; Ryant and Kipper, 2004) framework to cope with all verb alternations except those in which PP adjuncts are involved or for which no XTAG representation exists. The coverage of the other linguistic-based systems discussed above is on the other hand limited to a restricted set of alternations: light verb constructions (Fujita, 2005) or and synonymy and nominalisations (Brun and Hagège, 2003; Hagège and Roux, 2003). However, the approach proposed in this thesis can be extended to handle all kinds of verb alternations being based on an incremental framework and does not present any problem with alternations in which PP adjuncts occur.

Linguistic-based systems are often limited by the lack of lexical resources available, Brun and Hagège (2003); Hagège and Roux (2003) and Fujita (2005), for instance, use

hand-written rules to collect paraphrastic patterns, while in the Ryant and Kipper (2004) approach and in mine paraphrastic patterns are extracted directly from VerbNet, I then extract additional patterns by also using WordNet.
Finally, whereas none of the symbolic systems presented in this chapter has been evaluated, the system proposed in this thesis has been run against a test suite of examples extracted from VerbNet and the results show that it is very promising, namely it can recognise 86% of the sentence pairs as paraphrases if ontological knowledge is provided.

The symbolic approach to paraphrasing presented in this thesis is a first step towards a robust symbolic treatment of paraphrases. In fact, by integrating linguistic information extracted from existing large scale lexical resources such as VerbNet and WordNet with a robust parser like XIP, the proposed approach allows wide coverage extraction and recognition of paraphrastic patterns for verbs.
To improve coverage, however, much remains to be done. The approach needs to be extended to non canonical variants. Indeed, the evaluation is currently restricted to canonical alternation variants, that is, sentences without extraposition or pronominalisation, for instance. Because the basic XIP grammar produces functional dependencies (subject, object, etc.) and because the thematic rules used for semantic construction rely in part on such dependencies, it is likely that the treatment of alternations presented here straightforwardly extend to non-canonical variants.
Other needed extensions concern the treatment of other types of paraphrases such as those produced using intercategorial synonymy, morphoderivational variants (77a), nominalisation (77b), converse constructions (77c) and antonyms (77d).

(77) a. *John stopped smoking/Jean no longer smokes*

b. *The cost of the cruise is high/The cruise costs a lot*

c. *John lent a book to Marie / Marie borrowed a book from John*

d. *John is slow/John is not fast*

For morphoderivational variants, linguistic resources such as Celex[20] exists which could be integrated into XIP in a manner similar to the integration of VerbNet information. For converses and intercategorial synonymy, it is less clear, where the linguistic resources would come from.

[20] HTTP://WWW.RU.NL/CELEX/

Part II

Adjective-Based Entailment

Adjectives: Linguistic Background

A general problem of linguistic-based NLP approaches to textual entailment recognition is the lack of detailed resources providing the required knowledge and in particular, the lexical information these systems need to resolve inference.
As I discussed in the first part of this thesis, the organisation of linguistic information in semantic classes which group together items sharing common features such as proposed by Levin (1993) for English verbs, results in a valuable strategy for encoding lexical knowledge for computational applications. To show how such a strategy can enhance the performance of NLP systems to handle inference, I presented a robust linguistic-based approach to verb paraphrases recognition which uses the lexical knowledge encoded in Verbnet, an on-line wide coverage resource based on Levin's classification.

In the second part of this thesis, my aim is to extend Levin's approach to English adjectives, thereby providing a fine-grained semantic classification supporting the creation of lexical resources for computational systems which support adjective-based entailment recognition.

In linguistics, many theoretical attempts have been made to classify adjectives. Some focus on the analysis of their syntactic properties (see Huddleston, 1984; Quirk et al., 1985; Vendler, 1963, 1968), others focus on the analysis of their model theoretic properties (see Kamp, 1975; Keenan and Faltz, 1985; Keenan, 1987; Chierchia and Connell-Ginet, 1990; Kamp and Partee, 1995), and yet others concentrate on ontological distinctions (see Dixon, 1982, 1991; Frawley, 1992; Aarts, 1976).

In this chapter, I give an overview of these three types of classification concentrating on their most prominent representatives, namely (Huddleston, 1984), (Quirk et al., 1985), (Vendler, 1963) and (Vendler, 1968) for syntactic classifications; (Kamp, 1975), (Keenan and Faltz, 1985), (Keenan, 1987), (Chierchia and Connell-Ginet, 1990) and (Kamp and Partee, 1995) for semantic ones and (Dixon, 1982), (Dixon, 1991), (Frawley, 1992) and (Aarts, 1976) for ontological approaches.

6.1 Adjective Typology

As (Dixon, 1991) pointed out, while all languages provide means of modifying the meaning of nouns, they may differ in the syntactic form that such modification takes. Thus, some languages do not have adjectives and use verbs or nouns instead, to express noun modification. In English however, adjectives are the most important syntactic category associated with noun modification. As we shall now see, this category is not homogeneous and there is a number of properties with respect to which two adjectives may differ.
There are distributional differences in that an adjective can be used either in a predicative construction (78a) or as a noun modifier (78b).

(78) a. *The table is red.*

b. *This is a red table.*

Adjectives, like nouns, may or not subcategorise for complements. Thus for instance, while the adjective *loath* requires a sentential argument (e.g. (79)), *proud* takes an optional nominal complement and *brown* never admits of a complement.

(79) a. *Mary is loath to admit a mistake.*

b. **Mary is loath.*

(80) a. *Mary is proud of her son.*

b. *Mary is proud.*

(81) *This table is brown.*

Some adjectives have comparative and superlative forms, others not.

(82) a. *This painting is nicer that that one. / This painting is the nicest.*

b. **This table is more rectangular than that desk./ *This table is the most rectangular*

Semantically, different adjectives may behave differently with respect to entailment. For instance, while *white* licences the entailment in (83a), *alleged* does not (83b).

(83) a. *John is a white president* $\rightarrow$ *John is white.*

b. *John is an alleged murderer* $\rightarrow$ *John is a murderer.*

Further, adjectives may behave differently with respect to lexical relations and in particular, w.r.t., antonymy. For instance, while *dry* stands in a binary antonymy relation with *wet*, *round* and *rectangular* stand in a multiple opposition relation thereby licensing different inferences.

(84) a. *The cloth is dry $\rightarrow$ The cloth is not wet.*

b. *The cloth is not dry $\rightarrow$ The cloth is wet.*

c. *The table is rectangular $\rightarrow$ The table is not round.*

d. *The table is not rectangular $\not\rightarrow$ The table is round .*

In linguistic theory, many attempts have been made to classify adjectives with respect to their properties and in particular with respect to their syntactic, their semantic or their ontological properties. In what follows, I summarise each of these three main types of different classifications. I then situate the work presented in this thesis with respect to them and explain in what way, the classification proposed here differ from the existing proposals discussed in the chapter.

6.2 Syntactic Analysis of Adjectives

Syntax-based classifications of adjectives aim at individuating the set of syntactic properties which are representative of adjectival behaviour. I will here review the most prominent representatives of the syntactic approach namely, (Huddleston, 1984; Quirk et al., 1985; Vendler, 1963, 1968). Because Vendler's classification uses a set of criteria that encompasses the other two approaches, I will start by presenting (Huddleston, 1984; Quirk et al., 1985).

6.2.1 Huddleston Approach

Huddleston (1984) classifies adjectives by considering the different positions they can assume in a sentence and whether they can be used in the comparative and superlative form. More precisely, Huddleston uses the following selectional criteria for his classification:

1. Predicative Usage. The adjective co-occurs with a copular verb such as *be, seem, taste, smell, consider*.
   As the examples (85) show, not all adjectives can be used predicatively.

   (85) a. *This man is asleep.*

   b. *This man seems tired.*

   c. *This soup tastes delicious.*

   d. **This candidate is potential.*

2. Attributive Usage. The adjective modifies a noun.
   Again, as sentence (86a) shows, not all adjectives can be used attributively.

   (86) a. **This is an asleep man.*

b. *This is a tired man.*

c. *This is a delicious soup.*

d. *This is a potential candidate.*

3. Post- and pre-nominal usage. In the attributive usage, different adjectives may or not be placed after/before the nous.

(87) a. *This is the president elect/*This is the elect president*

b. *This is the former president/*This is the president former*

4. Gradability. The adjective has a comparative/superlative form and can be used with adverbs like *very*, *completely*, *extremely* which function as intensifiers. As illustrated in (88c), some adjectives are not gradable.

(88) a. *This book is more interesting than the film you are watching/This book is very interesting.*

b. **This table is more rectangular than the desk./*This table is very rectangular.*

On the basis of these properties, Huddleston individuates the five adjectival classes shown in Figure 6.1[21].

6.2.2 Quirk Approach

Quirk et al. (1985) use the following syntactic features to classify adjectives:

- syntactic function (attributive/predicative position)
- ability to be used in the comparative form
- ability to be modified by the intensifier 'very'

Although the adjectival features used for classification are very similar to those used by Huddleston (Quirk does not consider the postnominal use of adjectives), Quirk individuates the 6 adjectival classes shown in Figure 6.2 instead of the 5 proposed by Huddleston. Quirk divides adjectives into two main classes: the *central* and the *peripheral* class. The adjectives in the *central* class present all the characteristic features of adjectives: they accept both predicative and attributive usage and can be subdivided into two subclasses depending on whether they are gradable (*hungry*) or not (*infinite*).

[21]Adjectives are polysemous. Hence a given adjective can have conflicting properties under different meanings. For instance, *certain* has two main meanings namely, "definite but not specified or identified" and "destined or inevitable". Whilst in the first meaning, it accepts only an attributive usage (*"to a certain degree"; "certain breeds do not make good pets"; "certain members have not paid their dues"; "a certain popular teacher"; "a certain Mrs. Jones"*), in the second it accepts both an attributive and a predicative one (*"His fate is certain";"He faces certain death"*).

Central Adjectives

H_{A1} are the most frequent kind of adjectives, namely those which can be used predicatively, attributively and have comparative and superlative forms (e.g. *red, big, large*)

Non-central Adjectives

H_{A2} adjectives which are not gradable (e.g. *philatelic, male, pregnant*)

H_{A3} adjectives which cannot be used attributively (e.g. *afraid, asleep, glad*)

H_{A4} adjectives which have only attributive usage (e.g. *main, alleged, criminal, atomic*)

H_{A5} adjectives which are only used postnominally (e.g. *president elect* or *bishop designate*)

Figure 6.1: Huddleston Syntactic Classification of Adjectives

Peripheral adjectives are those whose properties more or less deviate from the standard ones. *Peripheral* adjectives, for example, can be either used predicatively (*afraid, asleep*) or attributively (*old, utter*). Moreover, the adjectives in each peripheral subclass can be further distinguished by considering the ability of taking comparative forms. For instance, both peripheral adjectives *afraid, asleep* have only predicative use, but while *afraid* is gradable, *asleep* is not[22].

		Attributive	Predicative	Comparison and Intensifier
Central	*hungry*	+	+	+
	infinite	+	+	-
Peripheral	*old*	+	-	+
	afraid	-	+	+
	utter	+	-	-
	asleep	-	+	-

Figure 6.2: Quirk Syntactic Classification of Adjectives

6.2.3 Vendler's Classification

Vendler's classification (Vendler, 1963, 1968) is also based on syntax but differs from the previous approaches in that it classifies adjectives in 8 classes by considering a wider set of typical adjectival alternations, i.e. paraphrastic syntactic constructions.

[22]Huddleston groups these two adjectives in the same class.

A1 is the class containing adjectives which allow both the `A N/N is A` patterns, i.e. both predicative and attributive patterns.

(89) a. *This is a rectangular table.*

b. *The table is rectangular.*

A2 adjectives allow the patterns `A N/A for an N`.

(90) a. *Dumbo is a small elephant.*

b. *Dumbo is small for an elephant.*

c. **This is rectangular for a table.*

A3 adjectives allow the patterns `A N/[V`$_N$`] Adv`$_A$, where Adv_A is morphologically related to A and V_N to N.

(91) a. *John is a beautiful dancer.*

b. *John dances beautifully.*

A4 is the class of adjectives which allow the patterns `A N/N is A to V`$_{imp}$, where V_{imp} is an action implied by the adjective and taking the modified noun as role.

(92) a. *This is a comfortable chair.*

b. *This chair is comfortable to sit on.*

A5 is the class of adjectives such as *clever, stupid, reasonable, kind* which allow the following patterns: `It is A of N SC/N is A SC`. This class of adjectives allows subject embedding (i.e. the subject can be realised as a clausal complement SC), and for which the preposition accompanying the nominal complement is *of*. This class corresponds to the *S-ofNP* adjective class described in Arnold (1989). As shown in (93c), the adjective describes a property (*stupid*) of the subject (*John*) of the clausal complement.

(93) a. *To take this job is stupid of John.* clausal-subject

b. *It is stupid of John to take this job.* it-extraposition

c. *John is stupid to take this job.*

A6 is the class of adjectives such as *possible, impossible* which allow the following patterns: `It is A for N SC/SC is A for N`. The adjectives in this class allow subject embedding but cannot be used attributively. They use the preposition *for* before the nominal complement and correspond to the *S-forNP* class (Arnold, 1989). As shown in (94c), adjectives in this class cannot be predicated of the subject (*you*) of the clausal complement, they can only be predicated of the nominalised sentences (SC, *that you work*) but with respect to the subject (N) of that sentence.

(94) a. *That you work is possible.* clausal-subject

b. *It is possible for you to work.* it-extraposition

c. **It is possible for me that you work.*

d. *That I succeed is impossible for you.*

A7 is the class of adjectives such as *useful, profitable, pleasant, necessary* that like the adjectives in the A_6 class allow subject embedding and use the preposition *for* before the nominal complement and denote properties which cannot be predicated of the subject of the clausal complement. But contrary to the adjectives in the class A_6, they can be predicated of the nominalised sentences (*that you work*), both with respect to the subject of that sentence, e.g. *you* in (95a), or with respect to another individual, e.g. *me* as in (95b) and (95c).

(95) a. *It is useful that you work.*

b. *That you work is useful for me.* clausal-subject

c. *Is is useful for me that you work.* it-extraposition

A8 is the class of adjectives such as *true, false, probable, certain, likely, unlikely* which allow the patterns `It is A SC/SC is A`. These adjectives can be used to predicate of a clausal subject but neither with respect to the subject of the sentence nor with respect to any another individual.

(96) a. *It is certain that he went away.*

b. *It is unlikely that he went away.*

c. **It is certain for me that he went away.*

d. **That he went away is unlikely for him.*

6.3 Semantic Analysis of Adjectives

In this section, I present some of the most prominent attempts to classify adjectives semantically, i.e. by focusing on their inferential properties.

6.3.1 Quirk Approach

Quirk et al. (1985) propose the following semantic features in order to characterise adjectival behaviour.

1. Gradability[23]
   An adjective is defined as *gradable* if it has comparative and superlative forms or if it can be modified by intensifiers like *very*, *extremely*.

[23] Gradability is considered a syntactic feature by Huddleston and a semantic one by Quirk. This illustrates the difficulty to characterise phenomena such as gradability, which are semantic in nature but have syntactic implications.

2. Staticity/Dynamicity
   Static adjectives such as *red* denote, like substantives, properties which are stable during time. *Dynamic* adjectives, e.g. *cruel*, behave like verbs denoting properties which can change during time. The notions of staticity/dynamicity are also known in the literature as *individual-level* and *stage-level* predicates (Carlson, 1977; Kratzer, 1995).
   Quirk et al. (1985) use the following syntactic constructions as tests to decide whether an adjective is static or dynamic[24]:

   C1: only dynamic adjectives can be combined with progressive forms of the verb *to be*

   (97) a. He is being cruel.
   *b. *He is being fat.*

   C2: only dynamic adjectives can be used in imperative constructions. Static adjectives do not occur in such patterns.

   (98) a. Be cruel!
   *b. *Be fat!*

3. Inherence/Noninherence
   Inherent adjectives characterise the referent of the noun. In (99), for instance, it is the individual referenced by the noun *man* who has the property of being *old*:

   (99) a. He is an old man.

   By contrast, *noninherent* adjectives characterise a property derived from the noun, not its referent. In (100), for instance, it is the *friendship* that is *old*, not the referent of the noun *friend*:

   (100) a. He is an old friend.

Type	Subclass	Gradable	Inherent	Static	Example
Central	Static	+	+	+	*She is a brave woman*
	Dynamic	+	+	-	*She is being very brave*
Peripheral	Static	+	-	+	*He is a firm[25] friend*
	Dynamic	+	-	-	*The actor is being wooden tonight*
	Non-gradable	-	-	+	*He is a medical student*

Figure 6.3: Quirk's Semantic Classification of Adjectives

[24] As we shall see in Chapter 7, additional tests can be used to distinguish between static and dynamic properties. In particular, Carlson (1977) defines 4 such tests.

By applying these criteria, Quirk et al. (1985) obtain the classification shown in Figure 6.3. The first class of adjectives, the class of *central* adjectives, contains adjectival items such as *brave*, which can be used both to express static or dynamic properties and are always inherent and gradable. The second class, the *peripheral* adjectives, contains adjectival items such as *firm, wooden, medical*, which are always noninherent. This class can be subdivided into two classes with respect to gradability: the gradable peripheral adjectives such as *firm, wooden* that can be static or dynamic, and the non-gradable peripheral adjectives such as *medical* that are always static.

6.3.2 Keenan Approach

Keenan and Faltz (1985) and Keenan (1987) individuate different semantic classes of adjectives by considering the following three semantic features: (i) whether the adjective is *restricting*, (ii) whether the adjective is *transparent* and (iii) whether the adjective is *absolute*.

Restricting (or affirmative) Adjectives

Members of this class are adjectives such as *red, round, tall, good, atomic*.
That an adjective is restricting means that for any N the denotation of the combination A N is a subset of the denotation of N:

(101) $\forall N : \| A\ N \| \subseteq \| N \|$

In natural language this describes the property that "adj N's are N". For example, *an atomic scientist* is a *scientist*, *a round table* is a *table* and a *skillful surgeon* is a *surgeon*. Not all adjectives are restricting; for instance, a *fake gun* is not a *gun*. Such non-restricting adjectives can be subdivided into two classes:

Negative

This class includes adjectives such as *fake, false* which assert the falsehood of the property expressed by the modified noun (e.g. *false diamonds* are not diamonds).

Conjectural

To this class belong adjectives such as *alleged, apparent* which question the truth of the property expressed by the noun they modify (e.g. *apparent victims* are not necessarily victims).

Transparent (extensional) Adjectives

This class groups adjectives such as *tall, fat, cold* for which the following holds:

[25] The adjective *firm* is used here in the sense of *loyal*, corresponding to the WordNet sense 11. So that in the sentence *firm* refers to the loyalty of the friendship rather than to the firmness/resoluteness of the friend.

if N_1 has the same denotation or extension as N_2 then A is transparent iff the denotation of the property expressed by combining A with N_1 is identical with the denotation of the combination of A with N_2, i.e.

$$A \text{ is transparent iff}$$
$$\forall N_1, N_2 : \| N_1 \| \equiv \| N_2 \| \longrightarrow \| \text{A N}_1 \| \equiv \| \text{A N}_2 \|$$

Therefore, if A is transparent, then the following entailments hold:

(102) $A\ N_1 \rightarrow A\ N_2 \wedge A\ N_2 \rightarrow A\ N_1$.

For example, if the property of being a *mother* exactly corresponds to that of being a *wife* (imagine a world in which all *mothers* are *wives*), then the extension of the intersection of *fat* with *mother* should be equal to the extension of the intersection of *fat* with *wife*.

(103) a. Mary is a fat mother → Mary is a fat wife

b. Mary is a fat wife → Mary is a fat mother

Not all adjectives are transparent; consider for example, adjectives such as *cruel, good, understanding* for which A $N_1 \not\rightarrow$ A N_2.

(104) a. Mary is a cruel mother ↛ Mary is a cruel wife

The fact that Mary is cruel as a mother does not imply that she is cruel in general, e.g. also as a wife.
Keenan claims that restricting adjectives can be transparent or not, while non-restricting adjectives are necessarily non-transparent.

Absolute (predicative) Adjectives

To this subclass belong adjectives such as *male*, *rectangular*. Such adjectives are categorical, i.e. they define a category, an absolute property. On the contrary, adjectives such as *tall, big* are scalar, and define a property which is not absolute but context-dependent.

(105) a. X is a big mouse → X is big with respect to the average size of a mouse

b. X is a big mouse ↛ X is big in general.

6.3.3 Chierchia Approach

In the classification of adjectives proposed in Chierchia and Connell-Ginet (1990) two semantic properties are considered as fundamental for describing adjectival behaviour: *extensionality* and *intentionality*. Thus, the class of *extensional* adjectives includes adjectives (e.g. *red, big, male*) which denote functions from individuals to individuals. This class can be subdivided into two subclasses: *intersective* (e.g. *red, male*) and *subsective* (e.g. *big, large, deep*) adjectives, depending on whether the [A N]$_{AP}$ phrase allows the inference of the property denoted by the adjective (i.e. $\parallel$ A N $\parallel \models \parallel$A$\parallel$) or not. As shown in (111), *subsective* adjectives (e.g. *big*) do not allow the inference of the A property as they lack an interpretation in absence of a standard of comparison.

(106) a. *Bixi is a big mouse.* $\rightarrow$ *Bixi is a mouse.*

b. *Bixi is a big mouse.* $\rightarrow$ *Bixi is big for a mouse.*

c. *Bixi is a big mouse.* $\not\rightarrow$ *Bixi is big.*

The class of *intensional* adjectives includes adjectives such as *past, former, fake, alleged*. These adjectives are property-modifying and can be interpreted as operators, i.e. they denote functions from properties to properties.

(107) a. *Victor is a former Catholic.* $\not\rightarrow$ **Victor is former.*

b. *Victor is a former Catholic.* $\not\rightarrow$ *Victor is a Catholic.*

c. *Victor is a former Catholic.* $\rightarrow$ *Victor was a Catholic.*

6.3.4 Kamp and Partee Approach

Perhaps the most complete and clear semantic classification of adjectives is that first proposed in Kamp (1975) and later reformulated in Kamp and Partee (1995). Kamp and Partee use as classification criteria the inference patterns shown by attributive adjective-noun combinations and individuate the following four semantic classes for adjectives.

Intersective Adjectives This class of adjectives includes adjectives such as *red, rectangular, male*. Adjective-noun combinations [A N]$_{AP}$ of members of this class entail the conjunction of the properties expressed by the noun N and by the adjective A.

(108) a. *This is a red car.*

b. $\rightarrow$ *This is red.*

c. $\rightarrow$ *This is a car.*

This behaviour can be described with the following two inference patterns:

(109) $A\ N \rightarrow A$
$A\ N \rightarrow N$

In set theoretic terms, the semantics of intersective adjectives correspond to the formula:

(110) $\parallel A \cap N \parallel \equiv \parallel A \parallel \cap \parallel N \parallel$

Subsective Adjectives This class of adjectives includes adjectives such as *big, good, old, atomic*. In this case, the $[\text{A N}]_{AP}$ phrase only entails the property expressed by the noun N.

(111) a. John is a good cook.

b. $\rightarrow$ John is a cook.

c. $\not\rightarrow$ John is good.

Thus, subsective adjectives allow only the inference pattern below:

(112) $A\ N \rightarrow N$

In set theoretic terms, the semantics of such adjectives are described by the formula:

(113) $\parallel A \cap N \parallel \subseteq \parallel N \parallel$

Privative Adjectives This class of adjectives includes adjectives such as *former, fake, past, fictitious*. Members of this class form $[\text{A N}]_{AP}$ phrases which entail the negation of the property expressed by the noun N.

(114) a. This is a fake fur.

b. $\not\rightarrow$ This is a fur

c. $\rightarrow$ This is not a fur

So that the following inference pattern is observed:

(115) $A\ N \rightarrow \neg\ N$

In set theoretic terms, the semantics of privative adjectives correspond to the formula:

(116) $\parallel A \cap N \parallel \not\subseteq \parallel N \parallel$ *or equivalently* $\parallel A \cap N \parallel \cap \parallel N \parallel \equiv \emptyset$

Plain Nonsubsective Adjectives This class includes adjectives such as *alleged, putative*. $[\text{A N}]_{AP}$ phrases formed with members of this class entail neither the A nor the N property.

(117) a. This is an alleged murderer.

b. $\not\rightarrow$ this individual is a murderer

c. $\not\rightarrow$ this individual is not a murderer

This behaviour can be described by the inference pattern:

(118) $A\ N \rightarrow N \vee \neg\ N \rightarrow$ *true*

In set theoretic terms, the semantics of such adjectives can be expressed by the formula:

(119) $\| A \cap N \| \subseteq \mathcal{D}$

where $\mathcal{D}$ is the domain of discourse.

It is worth noting that several criteria proposed by different researchers in fact describe the same or similar properties. For example, the notion of *inherence* described in Quirk et al. (1985) corresponds to the notion of *absoluteness* or *predicativeness* of Keenan (1987). Similarly, the inference patterns used as classification criteria by Kamp (1975); Kamp and Partee (1995) partially correspond to those used by Keenan and Faltz (1985); Keenan (1987). Thus, Kennan's *restricting adjectives* correspond to Kamp and Partee's *subsective* adjectives while his *absolute adjectives* are in fact Kamp and Partee's *intersective* class.

6.4 Taxonomies of Adjectives

Dixon (1982) was one of the first authors to propose a taxonomy of adjectives. This taxonomy originates from his study on the distribution of adjectival meaning in different languages. By analysing 20 languages from different language families, he found a distinction between (i) languages which make more or less use of adjectives and (ii) languages such as Yurok[26] which have no adjectives at all and use nouns or verbs to express the meaning that in other languages is expressed by means of adjectives. As result of this study, Dixon comes up with the observation that in the languages which have adjectives, the meaning expressed by the adjectives is quite constant and can be expressed by the following seven universal properties or semantic categories:

1. DIMENSION
2. COLOUR
3. VALUE
4. AGE
5. PHYSICAL PROPERTY
6. HUMAN PROPENSITY
7. SPEED

Later, Dixon (1991) proposes a more detailed taxonomy of adjectives based on the ten categories shown in Figure 6.4.
Similarly, Aarts (1976) defines three major semantic categories for adjectives: (i) static, (ii) physical, and (iii) dimensional.

Dimensional
horizontal, vertical, quantity, general, size, time, duration, frequency, iteration

[26]Yurok is an Algonquian language spoken in California

1- `Dimension`	*big, large, little, small, ...*	
2- `Physical property`	*hard, soft, heavy, light, ...*	
3- `Speed`	*fast, quick, slow, ...*	
4- `Age`	*new, young, old, ...*	
5- `Colour`	*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- `Fond`	*fond*
	9.2- `Angry`	*jealous, angry, ...*
	9.3- `Happy`	*anxious, happy, ...*
	9.4- `Unsure`	*certain*
	9.5- `Eager`	*eager, ready*
	9.6- `Clever`	*clever, stupid, generous*
10- `Similarity`	*similar, different, ...*	

Figure 6.4: Dixon Taxonomy of Adjectives

Non-Dimensional
: substance, solidity, liquidity, gaseousness, texture, luminosity, humidity, temperature, colour, weight, smell, taste, vision, touch, sound, musical sound, weather, fixity, property, content, corp. condition (*hungry*), activity, corp. function (*blind*), velocity

Non-Physical
: emotion, attitude, intellect, truth, communication, manner, evaluation, degree, modality

And Frawley (1992) extends the classification categories of Dixon by defining subclasses and obtains the taxonomy shown in Figure 6.5.

6.5 Discussion

The linguistic classifications of adjectives presented in this chapter provide a good starting point for building a classification that supports adjective based inference. Indeed, all

1- Dimension		
2- Physical property		
	2.1- Sense	
	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- Colour	*red, blue, black, ...*	
...		
9- Human Propensity		
	9.1- Mental State	*jealous, happy, loyal, ...*
	9.2- Physical State	*weak, sore, thirsty, ...*
	9.3- Behaviour	*wild, argumentative, funny, ...*
...		

Figure 6.5: Frawley Taxonomy of Adjectives

aspects (syntactic, semantic and ontological) of these classifications are relevant to the inference problem.

Trivially, syntax is necessary, as a scaffold for semantic construction. In order to define the mapping between lexical semantics, syntax and the meaning of phrases, it is necessary to know which adjectives license which syntactic constructions. Furthermore, as shown in the first part of this thesis, syntax is also useful for detecting semantic equivalences between distinct syntactic patterns such as Vendler's alternations.

(120) a. *John is stupid to take this job*

b. *It is stupid of John to take this job*

Semantics and in particular, the type of model theoretic classification proposed by Kamp and Partee, is trivially relevant as it directly specifies adjective based inference patterns.

Finally, ontological classifications permit structuring the set of adjectives into taxonomies that should be useful for modelling knowledge based reasoning.

Nonetheless, although they are undeniably relevant for defining a classification that help predict adjective based inferences, these linguistic classifications alone do not suffice. This is because they each concentrate on a single dimension (syntax, semantics or ontological knowledge) and as a result, fail to predict interactions between these dimensions which impact permissible inferences.

There are clear interactions between syntax and semantics. For instance, although adjectives such as *red* and *large* are grouped together in the syntactic analysis proposed by Huddleston and Quirk in the class of *central* adjectives (they participate in the attributive/predicative alternation and are gradable), they display different semantic properties and would be distinguished in Kamp and Partee's classification as belonging to the *intersective* and *subsective* classes respectively.

Conversely, while Kamp and Partee gather adjectives like *small*, *stupid* and *probable* in the subsective class, these same adjectives are split by Vendler syntactic analysis into three different classes A1 (e.g. *small*), A5 (*stupid*) and A6 (*probable*). Interestingly, each of theses classes corresponds to three distinct taxonomical classes in Dixon's classification namely, *dimensional* (*small*), *value* (*stupid*) and *qualification* (*probable*).

There are also idiosyncratic interactions between model theoretic semantics and syntax. For instance, none of the proposed classification can account for the difference in entailment behaviour between (121a) and (121b).

(121) a. *Bibi is a tall mouse* $\not\rightarrow$ *Bibi is tall*

b. *Bibi is a 30 cm tall mouse* $\rightarrow$ *Bibi is 30 cm tall*

In both cases, the same adjective (*tall*) is involved but the adjective occurs in a different syntactic context. As a result, in (121a), *tall* behaves as a subsective adjective (the entailment is not licensed) whereas in (121b), the additional measure modifier *30 cm* induces an intersective reading for the overall modifier *30 cm tall*. Given the Kamp and Partee classification alone, the inference in (121b) will not be predicted.

Further, interactions with lexical relations are not taken into account by any of the proposed classification although again, such interactions clearly impact inference. For instance, although both *dry* and *rectangular* can be semantically classified as *intersective*, they present different inference patterns with respect to antonymy since *dry* enters into a binary antonymic relation with *wet* whilst *rectangular* does not have a unique antonym and enters into multi-opposition relations with a set of antonyms. As a result, both *dry* and *rectangular* license the inference pattern N IS A $\rightarrow$ N IS NOT ANTONYM(A) but only *dry* licenses the additional entailment illustrated in (b) namely, N IS NOT A $\rightarrow$ N IS ANTONYM(A).

(122) a. *The dishcloth is dry* $\rightarrow$ *The dishcloth is not wet*

b. *The dishcloth is not dry* $\rightarrow$ *The dishcloth is wet*

(123) a. *This table is rectangular* $\rightarrow$ *This table is not round*

b. *This table is not rectangular* $\not\rightarrow$ *This table is round*

Similarly, derivational morphology needs to be taken into account so as to capture semantic equivalences based on derivational links between adjectives and nouns (124), verbs (125) and adverbs (126).

(124) a. *Wikipedia indicates that the height of Mount Everest is 8,849 meters* $\leftrightarrow$ *Wikipedia indicates that Mount Everest is 8,849 meters high*

b. *That John is present to the party offends Mary* $\leftrightarrow$ *John's presence to the party offends Mary*

(125) a. *The boy is asleep* $\leftrightarrow$ *The boy is sleeping*

b. *The boat is afloat* $\leftrightarrow$ *The boat is floating*

(126) a. *Mary is doing a careful survey of the flat* $\leftrightarrow$ *Mary is carefully surveying the flat*

b. *John's sleep is sound* $\leftrightarrow$ *John is sleeping soundly*

In short, to develop a classification of adjectives that adequately predicts their inferential behaviour, it is necessary to consider not only, each of the dimensions considered by the three main types of classification (syntactic, semantic, ontological) discussed in this chapter but also their interaction. Furthermore, these classifications need to be extended with information about lexical relations (e.g., antonymy and synonymy) and about derivational morphology relations. In the following chapters, I present a classification that aims to fulfil these criteria.

6.6 Summary

In this chapter, I reviewed several approaches to adjective classification, some based on analysis of syntactic features, some which focus on model theoretic properties and some based on ontological distinctions.
In the next chapter, I show that by merging together the syntactic and semantic criteria used in the discussed classifications with lexical semantic and morphoderivational ones, a classification can be defined which correctly predicts a wider range of adjective-based inferential patterns.

A Classification of English Adjectives

As shown in the previous chapter, previous classifications of adjectives are based on a single type of knowledge such as syntax or semantics. As a result, they fail to consider the interplay between different levels of linguistic description although as I have argued, syntax, semantics, lexical semantics and derivational morphology are all relevant factors in determining possible inferences. These approaches therefore, are limited in their ability to handle adjective-based inferences.

In this chapter, I propose a classification that builds on Levin's (1993) idea that the semantic properties of the members of a given syntactic category (e.g., verbs) are reflected in the syntactic patterns (and in particular, in so-called syntactic alternations) in which they participate. The proposed classification departs from Levin's proposal however, in that it uses a wider set of classification criteria than just alternations. The motivations for this are the following.

First, this permits taking into account the various factors influencing the inferential behaviour of adjectives i.e., syntax but also model theoretic semantics, lexical semantics and derivational morphology. More specifically, the proposed classification aims to integrate the various syntactic, model theoretic and ontological constraints identified by the literature on classification and to further extend it with information from lexical semantics and derivational morphology.

Second, the high number of classification criteria put to use helps in defining finer grained, homogeneous classes. This should in particular, help avoiding a problem that was repeatedly pointed out with Levin's classification namely, that some verbs were members of classes defined by contradictory alternations (such verbs were then predicted to be both p and $\neg p$) without there appearing to be a corresponding semantic ambiguity (in all classes, the verbs seem to take on the same meaning).

Third, considering the fine grained interactions between the different levels of linguistic knowledge provides important clues for specifying the compositional and decompositional semantics of the members of the different classes.

In sum, I propose a classification of adjectives which adapts Levin's classification proposal to adjectives and extends it by integrating further classification criteria such as model theoretic semantic, lexical relations and derivational morphology. Furthermore, this proposed classification associates each classification criterion with the inference pattern(s) it licenses. As a result, it can handle an extensive set of adjective based inferences and should provide a useful basis for modelling entailment recognition between sentences containing adjectives.

The chapter is structured as follows. I first describe how a sample set of adjectives was chosen (section 7.1) on which to base the classification. I then describe the set of selectional criteria used for the classification (section 7.2) and indicate how the adjectives in the sample set were tested for these criteria. Finally, I show how the systematic application of these criteria to the sample set of English adjectives leads to the definition of a set of fine grained semantic classes (section 7.3).

7.1 Sample Set of Adjectival Items

One feature of adjectives that makes their analysis and classification difficult is polysemy. Adjectives, like other categories, 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*.

(127) a. This bag is heavy

b. John is a heavy smoker

It is clear that *heavy* in (127a) has a dimensional meaning, while *heavy* in (127b) is a quality adjective. In order to cope with this problem, I rely on WordNet (Fellbaum, 1998)[27] and define an adjectival item as corresponding to a WordNet synset, i.e. to the reading corresponding to the interpretation of the adjective in the given example sentence. So for instance, the different interpretations of the adjective *heavy* in the sentences (127a) and (127b) can be mapped to the WordNet senses 1 and 14:

`heavy`$_{1}$ (vs. light): (of comparatively great physical weight, "a heavy load")
`heavy`$_{14}$(prenominal) (vs. temperate): (excessive in behaviour; "heavy investor")

Therefore, different WordNet readings of the same adjective can have different properties and can belong to different semantic classes.

In applying my classification, I will therefore start from a set of adjectival meanings (as identified by WordNet synsets), not adjectives. The mapping of adjectival items to WordNet senses is motivated by two important facts. First, it allows access to the linguistic knowledge about adjectives encoded in WordNet (e.g. antonyms, hyponyms, hyperonyms, related nouns and verbs, etc.); and second, as this lexical resource has a wide usage in NLP, it enables the classification to be used in real applications.

[27]WordNet (Gross and Miller, 1990) defines different readings of adjectives by means of antonymy: if an adjective has in two different contexts two different antonyms, then the adjective is assigned two different readings, one for each antonym.

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- Colour	*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, ...*

Figure 7.1: Taxonomy of Adjectives

The set of domain independent adjectival items on which the proposed classification will be applied was defined as follows.

First, Dixon (1991)'s work was used to define a basic set of adjectives representative of the main ontological classes identified by Dixon. The ontological classes used are given in Figure 7.1.

Second, for each ontological category, a set of items displaying different syntactic (e.g. predicative, attributive, postnominal use, clausal complement, etc.), semantic (e.g. intersective, subsective, privative, plain nonsubsective adjectives), lexical semantic and morphoderivational properties was singled out.

Third, the initial sample given by the preceding two steps, was further expanded with synonyms, similar words, hyponyms and antonyms.

To integrate the results provided by previous classifications, particular attention was paid to include in the created sample, the largest possible number of adjectival items considered by the classification discussed in chapter 6. The final sample includes about 500 adjectival items corresponding to about 450 WordNet synsets. These 500 items are listed in the index. Furthermore, for each item, the index references two page numbers, the first one refers to the page where an informal description of the item class is given and the second to the page in the appendix where a more detailed description of that class is given.

7.2 Classification criteria and methodology

The set of linguistic properties that have been used as selectional criteria to classify English adjectives were obtained by merging the different syntactic and semantic features considered in the classifications of Chapter 6 together with lexical semantic and morphoderivational criteria such as antonymy, nominalisations and verbalisations which were not considered in previous works.

In what follows, I first explain how criteria were tested for (section 7.2.1). I then present the criteria used together with the inference patterns they each license.

7.2.1 Testing methodology

To test for a given criteria, a number of methods can be used such as in particular, introspection, literature survey, corpus analysis and multiple native speakers questionnaires. On practical or on theoretical grounds however, all of these methods are not all always possible. Theoretically, corpus based verification of inference patterns for instance, is difficult and often impossible. Practically, time and/or financial means do not always permit the use of multiple native speakers questionnaires.

The methodology adopted here is a combination of literature survey, double judgements and corpus search. Multiple native speaker surveys and/or more sophisticated corpus analyses were not carried out mainly because of limited time and financial means although clearly, they would provide a better quality warrant for the proposed classification. It would in particular, be useful to have the proposed classes be cross examined by several native speakers.

Generally, the way criteria were tested for is as follows. From the literature survey on existing classifications, I took on the data attested by the linguists. Whenever the adjectives or the criteria under consideration had not been treated in the literature, I resorted either to manual Google searches for particular patterns or to a double judgement by myself and by a German student of English linguistics. In the sections that follow, I will make precise, how each criteria was tested for or in other words, which methodology was used to determine whether or not a given adjectival item satisfies a given criteria.

7.2.2 Model Theoretic Properties

In order to individuate a set of model theoretic properties of adjectives which can be used as selectional criteria for the classification of English adjectives, I rely on Kamp and Partee (1995) and Keenan (1987).
As discussed in Chapter 6, Kamp and Partee describe the semantics of adjective-noun combinations, i.e. the semantics of the attributive use of adjectives, by means of the inference patterns these combinations allow. In particular, the main criterion they use for classification is whether it is possible to infer from the $[\text{A N}]_{AP}$ phrase the denotation of the property expressed by the adjective A or the denotation of the property expressed by the noun N or both of them. In this way, they come up with four semantic classes, namely *intersective*, *subsective*, *privative* and *plain nonsubsective*. In sum:

- An adjective such as *red* is *intersective*, as the $[\text{A N}]_{AP}$ phrase textually entails[28] both N and A.

- An adjective such as *big* is *subsective* because the $[\text{A N}]_{AP}$ phrase only entails N.

- An adjective such as *fake* is *privative* because the $[\text{A N}]_{AP}$ phrase entails $\neg$ N.

- An adjective like *alleged* is *plain nonsubsective* because nothing can be inferred from the $[\text{A N}]_{AP}$.

However, this classification does not account for the fine differences in the meaning of adjectives. For instance, the different inferential behaviour of the adjectives *false* and *fictitious* cannot be predicted.

(128) a. *John is a fictitious friend → John is not a (real) friend*

b. *John is a false doctor → John is not a (genuine) doctor*

[28]Strictly speaking, a noun phrase cannot entail a noun or an adjective since all these items denote sets of individuals, not propositions. Hence at the NP level, we need to talk about set relations whereas at the sentential level, we can speak of entailment. To simplify matters, we will speak of textual entailment (or entailment written →) whenever either holds.

c. *John is a fictitious friend* $\rightarrow$ *John is fictitious (as an entity)*

d. *John is a false doctor* $\not\rightarrow$ *John is false (as an entity)*

The adjectives *false* and *fictitious* should both be classified as *privative* as they both (see (128a) and (128b)) entails the negation of the noun property but while *fictitious* denotes a property of the individual denoted by the noun (128c), *false* does not (128d).
In order to distinguish between these two different inferential behaviours I rely on the notion of *absoluteness* proposed by Keenan (1987).
Keenan uses the notion of *absoluteness* to describe adjectives which characterise absolute properties of the noun they modify.

(129) a. *This is a red table* $\rightarrow$ *This is red*

b. *John is a good cook* $\not\rightarrow$ *John is good*

c. *John is a good cook* $\rightarrow$ *John is good as a cook*

For instance, the adjective *red* in (129a) is *absolute* as it characterises a property of the individual referred to by *table*. By contrast, the adjective *good* in (129b-c) is *nonabsolute* as it characterises the property of *being a cook* and not John himself. Therefore, *absoluteness* characterises all adjectives which allow to infer from the $[\text{A N}]_{AP}$ phrase that A is true of the individuals denoted by N.
The notion of *absoluteness* explains the different inferential behaviour of adjectives such as *fictitious* and *false*: these adjectives in fact are both *privative*, but while *fictitious* is *absolute*, *false* is not. As a result, *fictitious* entails both $\neg$ N and A whereas *false* entails only $\neg$ N.

In order to further refine the model theoretic semantic differences between adjectives, I rely on Keenan and Faltz (1985); Keenan (1987) notion of *transparency*. As we saw in Chapter 6, *transparency* is defined in terms of set theory. Consider for instance (130), if the class of *wives* in a possible world corresponds to (is exactly the same as) the class of *mothers*, then an adjective like *tall* is transparent but an adjective like *cruel* is not because the property of *being cruel* is not shared between the class of mothers and that of wives although they are made up of the same individuals.

(130) a. *X is a tall mother* $\rightarrow$ *X is a tall wife*

b. *X is a cruel mother* $\not\rightarrow$ *X is a cruel wife*

That means that if an individual i is described by n properties, then the combination of a *nontransparent* adjective with one of these properties modifies the particular property and not the individual itself. So that in (130) the individual denoted by the noun *mother* is not *cruel* in general, it is *cruel* just in her role as a mother.

It is worth noting that if an adjective is *absolute* it is necessarily *transparent* but if it is not it can be both *transparent* and *nontransparent*. As an example, consider the adjectives *cruel* and *tall*. They are both *nonabsolute* as they both (131) do not allow the inference of the adjectival property, but the adjective *tall* is *transparent* whereas the adjective *cruel* is not.

(131) a. *X is a tall woman $\not\rightarrow$ X is tall (in general)*

b. *X is a cruel mother $\not\rightarrow$ X is cruel (in general)*

Transparency in fact, describes properties of individuals with respect to specific sets and not in general as *absoluteness*. So that, someone can be not *tall* in general but *tall* in a particular subset of the domain.
Thus, *transparency* can be used as a criterion to distinguish between subsective adjectives such as *cruel* which are *nonabsolute* and *nontransparent* and characterise the nouns they modify and subsective such as *big, tall, fat* which are *nonabsolute* but *transparent* meaning that they describe properties of individuals which are context dependent.

The following table summarises the inference patterns licensed by each of these properties.

Intersective	AN $\rightarrow$ A, N	*red*
Subsective	AN $\rightarrow$ N	*tall, cruel*
Privative	AN $\rightarrow$ $\neg$N	*fake*
Plain non subsective	No entailment	*alleged*
Absolute	AN $\rightarrow$ A	*fictitious*
Transparent	(N1 $\leftrightarrow$ N2) $\rightarrow$ (AN1 $\leftrightarrow$ AN2)	*tall,fictitious*

To classify adjectives with respect to each of these 6 properties, I relied mainly on the literature as corpus analysis did not seem usable (it was unclear how such semantic properties could be tested in corpus). For the main four properties (intersectivity, subsectivity, privativity and plain non subsectivity), the sample set was classified mainly based on Kamp and Partee's analysis while for the other two properties, Keenan's work was used. In the case of adjectives not classified by other authors, the judgement on whether a pattern holds or not, was based on the comparison of the judgements of two raters. In case of disagreement, the pattern have been disallowed for the given adjective. In practise, I found that such inference patterns, especially the privative one, are very difficult to judge, as human inference is a multimodal process in which other mechanisms than pure speech (e.g. sight, etc.) are involved.

The resulting classification comprises 6 classes corresponding to the following combinations of properties:

Intersective	AN $\rightarrow$ A, N	*red*
Subsective and transparent	AN $\rightarrow$ N (N1 $\leftrightarrow$ N2) $\rightarrow$ (AN1 $\leftrightarrow$ AN2)	*tall*
Subsective and non transparent	AN $\rightarrow$ N	*cruel*
Privative, absolute, transparent	AN $\rightarrow$ $\neg$N,A (N1 $\leftrightarrow$ N2) $\rightarrow$ (AN1 $\leftrightarrow$ AN2)	*fictitious*
Privative, non absolute, non transparent	AN $\rightarrow$ $\neg$N	*fake*
Plain non subsective	No entailment	*alleged*

Not all possible combinations of properties are observed. Obviously, neither subsective nor plain non subsective can be absolute adjectives (which entails A). In the first case, this would be contradictory with the fact that subsective adjectives do not entail A, in the second, this would contradict the fact that plain non subsective adjectives do not license any entailment. More interestingly, I noted that absolute adjectives are a subset of the transparent one. That is, I found no adjectives that were "intersective and non transparent", "privative, absolute and non transparent", "privative, non absolute and transparent". In other words, all intersective adjectives from the sample were found to be transparent, all absolute privative one were found to be transparent and all privative, non absolute were non transparent. Further, I observed that the transparent nonabsolute adjectives denote context dependent properties, i.e. are dimensional adjectives. Based on these observations, the 6 classes observed were named as follows:

Intersective	A N $\rightarrow$ A, N	*red*
Dimensional Subsective	A N $\rightarrow$ N	*tall*
	(N1 $\leftrightarrow$ N2) $\rightarrow$ (A N1 $\leftrightarrow$ A N2)	
Non dimensional Subsective	A N $\rightarrow$ N	*cruel*
Absolute Privative	A N $\rightarrow$ $\neg$N,A	*fictitious*
	(N1 $\leftrightarrow$ N2) $\rightarrow$ (A N1 $\leftrightarrow$ A N2)	
Non absolute Privative	A N $\rightarrow$ $\neg$N	*fake*
Plain non subsective	No entailment	*alleged*

7.2.3 Lexical Semantics

As lexical semantic selectional criteria, I consider (i) the type of the antonymy relation in which adjectives are involved, (ii) dynamicity/staticity and (iii) gradability.

Antonymy

Antonymy[29] seems to be the basic lexical semantic relation for many adjectives. As experiments on word association tests (Deese, 1964, 1965) show, when presented with an adjective, speakers naturally tend to associate it with its antonym.

As discussed in Cruse (1986), the term *antonymy* covers different kinds of opposite polarity relations between adjectives namely, binary opposition, contraries and multiple oppositions.

- **Binary oppositions** cover pairs such as *finite/infinite* which license the following inference pattern between two adjectives $A1$ and $A2$:

$$A1 \rightarrow \neg A2 \ \wedge \ \neg A2 \rightarrow A1$$

[29] I use antonymy as a distinct classification criterion although some authors (Lyons, 1977) claim that this test corresponds to gradability: contraries adjectives are gradable, binary (or contradictory) are not. But there are also adjectives such as *dry* which are both gradable and binary.

Example:
X is finite $\rightarrow$ *X is not infinite*
X is not infinite $\rightarrow$ *X is finite*

- **Contraries** are pairs such as *long/short* where the implication is unidirectional:

$$A1 \rightarrow \neg A2 \ \wedge \ \neg A1 \not\rightarrow A2$$

$$A2 \rightarrow \neg A1 \ \wedge \ \neg A2 \not\rightarrow A1$$

Example:
X is long $\rightarrow$ *X is not short*
X is not long $\not\rightarrow$ *X is short*

- **Multiple oppositions** involve a finite set of adjectives (e.g. *liquid/solid/gaseous*) which are pairwise mutually exclusive. For a set of adjectives $A_1 \ldots A_n$, standing in a multiple opposition relation, the following axiom schemes will be licensed:

$$\forall i, j \ s.t. \ 1 \leq i, j \leq N \quad and \ i \neq j :$$

$$A_i \rightarrow \neg A_j \quad and \quad \neg A_i \not\rightarrow A_j$$

Example:
X is liquid $\rightarrow$ *X is neither solid nor a gas*
X is not liquid $\not\rightarrow$ *X is either solid or a gas*

It should be noted that negated $[\text{A N}]_{AP}$ phrases such as (132) express ambiguous statements so that the antonymic inference patterns sketched above only hold for adjectives occurring in predicative position.

(132) a. *This is not a red table* $\not\rightarrow$ *This is not red*

b. *This is not a red table* $\not\rightarrow$ *This is not a table*

c. *This is not a red table* $\rightarrow$ *This is not red or This is not a table*

Testing for antonymy was done partly using WordNet and partly using the judgements of two annotators. Adjectives that had more than one antonym in WordNet, were assigned a multiple opposition relation. Adjectives with a single Wordnet antonym on the other hand, were manually tested to decide whether the antonymic relation was binary or contrary. In such cases, I either adopted the classification assigned by the literature (Cruse, 1986) or, if the adjective was not accounted for in previous work, two human raters decided whether the formula (133) is satisfiable.

(133) $X \in \neg A \wedge X \in \neg Anto(A)$

Thus, if (133) is satisfiable the adjective is assigned a contrary antonymic opposition relation, otherwise a binary opposition. In case of disagreement between the raters, the adjective was assigned the contrary opposition.

X is not finite and X is not infinite	*NotSatisfiable* $\rightarrow$*binary*
X is not small and X is not big	*Satisfiable* $\rightarrow$*contrary*

Staticity vs. Dynamicity

As a further test, I consider whether the property expressed by the adjective is stable over time or not. As we saw in Chapter 6, this property is related to the notions of *individual-level* vs. *stage-level* predicates defined in Carlson (1977); Kratzer (1995).

(134) a. *I'm sitting on this chair.* *Stage-level predicate/Dynamic*

b. *I have brown hair.* *Individual-level predicate/Static*

For instance, *I'm sitting on this chair* is a transitory property, a stage-level or dynamic property. By contrast, *I have brown hair* is a permanent or stable property, a property of an individual, not of a state. Staticity/dynamicity can be tested by means of the following criteria:

1. static predicates are odd with temporal modifiers (Carlson, 1977)

   (135) a. *John was drunk yesterday/last month/ a year ago*

   b. **John was intelligent yesterday/last month/ a year ago*

2. locative modification is quite impossible for static predicates (Carlson, 1977)

   (136) a. *John is always sick in France.*

   b. **John is intelligent in France.*

3. static predicates cannot occur as arguments of perception verbs like *see* (Carlson, 1977)

   (137) a. *I saw John drunk.*

   b. **I saw John intelligent.*

4. the coda position of `there` sentences does allow only dynamic adjectives (Carlson, 1977)

   (138) a. *There are two men drunk/sick/available.*

   b. **There are two men intelligent/white/altruistic.*

5. only dynamic predicates may become adverbs using -ly (Quirk et al., 1985):

   (139) a. *He smiled cruelly.*

   b. **He ate fatly.*

6. only dynamic predicates can be combined with progressive forms of the verb *to be* (Quirk et al., 1985):

   (140) a. *He is being cruel.*

   b. **He is being fat.*

7. only dynamic predicates can be used in imperative constructions (Quirk et al., 1985):

(141) a. *Be cruel!*

b. **Be fat!*

Quirk et al. (1985) claim that there is a link between the notion of stage-level/individual level predicates and transparency. They assert that all stage-level (dynamic) predicates are nontransparent. But as (142) show, this does not hold. Therefore, I use transparency and stability as two distinct test criteria.

(142) a. *cruel mother* $\wedge$ *mother=wife* $\not\rightarrow$ *cruel wife* s-level/dynamic, nontransparent

b. *drunk mother* $\wedge$ *mother=wife* $\rightarrow$ *drunk wife* s-level/dynamic, transparent

c. *red car* $\wedge$ *car=object* $\rightarrow$ *red object* i-level/static, transparent

Staticity was tested using corpus analysis by submitting queries to Google which follow the constructions used in examples (135,136,137,138,139,140,141). I also ranked the different tests by order of importance. Depending on the number of documents Google found to contain the query and on the rank of the corresponding query, the adjective was classified either as static or as dynamic. For example, Figure 7.5 shows the results of this analysis for the adjectives *stupid* and *intelligent* (the order gives the ranking of the tests):

Adjective	Query	Documents Found
stupid	'He is being stupid'	18,300
	'He was stupid yesterday'	4
	'He is always stupid'	4
	'I saw him stupid'	1
	'He smiled stupidly'	717
Result :	dynamic	
intelligent	'He is being intelligent'	8
	'He was intelligent yesterday'	0
	'He is always intelligent'	105
	'I saw him intelligent'	2
	'He smiled intelligently'	6
Result:	Static	

Figure 7.2: Google Results for Stability Patterns

The progressive test illustrated in (140) was given first rank. A high value from the Google search indicates that the adjective is dynamic and a low value that it is static. Hence *stupid* was classified as dynamic and *intelligent* as static.

Clearly the methodology used to test for dynamicity/staticity could be improved. As illustrated by the example, the constructed Google queries are very specific (words are used not regular expressions) thereby diminishing the number of results. Further, the various tests can deliver contradictory information (one Google test might suggest staticity and another dynamicity) so that the ranking of the various testing criteria becomes very important. A weighting scheme or a scheme based on optimality theory might be more relevant here which would permit combining the various scores in a general way.

Gradability

Gradability is a further criterion which permits differentiating between different adjectival types. As shown in (143), syntactically gradable adjectives differ from nongradable ones in that they are acceptable in comparative and superlative forms and/or can be modified by intensifiers (e.g. *very*, *quite*, *rather*) and sufficiency morphemes (e.g. *too, enough, so*)[30].

(143) *a.* *Chicago is larger than Rome.*

b. *The Mars Pathfinder mission was very expensive.*

c. *This rod is too bent to be used for this purpose.*

d. **This table is very rectangular.*

e. **This murderer is more alleged than that one.*

Gradable adjectives describe measurable properties so that the standard meaning assigned to such adjectives corresponds to the assignment of a value to a property along a given measuring scale (see Seuren, 1973; Bartsch and Vennemann, 1973; von Stechow, 1984; Heim, 1985; Bierwisch, 1989; Klein, 1991; Kennedy and McNally, 1999, 2005). For instance, the adjective *expensive* describes the value of the property COST, while the adjective *fast* describes the value of the property SPEED.

As shown by Kennedy and McNally (2005), gradable adjectives can furthermore be divided into subclasses depending on the typology of the scale describing the property they measure.

As Figure 7.3 shows, a scale can be *open* in that it has neither minimal nor maximal element. It can be *lower* or *upper closed* in that it has either a minimal or a maximal element. And it can be *closed* in that it has both a minimal and a maximal element. Accordingly, Kennedy (2004) distinguishes between adjectives such as *dry, asleep, full, empty* which have a closed or a partially closed scale and which are also called *absolute*[31]

[30]Nongradable adjectives express a property which cannot be intensified using degree adverbs such as *very*. When such adjectives are modified by degree adverbs, the effect is to give emphasis rather than to express the degree of the property expressed, e.g. *He is very male, That's very true*, *It was a very black day.*

[31]The term "absolute" used by Kennedy has a different meaning than the term "absolute" discussed above that was introduced by Keenan.

OPEN	○———————○
LOWER CLOSED	•———————○
UPPER CLOSED	○———————•
CLOSED	•———————•

Figure 7.3: Typology of scale structures

or *context insensitive* adjectives as their interpretation is not dependent on the context, and adjectives such as *tall, expensive, big, fast* which have open scales and are also called *relative* or *context sensitive* as their interpretation depends on the context, and is underspecified. Thus, the adjective *expensive* has an open scale and is context sensitive as its interpretation depends on the individuation of a standard of comparison which varies from context to context. For instance, in order to compute the interpretation of (144a), i.e. to decide whether the sentence is true or false one needs a standard of comparison.

(144) a. *The coffee is expensive in Rome*

b. *The dishcloth is dry*

A coffee which is expensive in Rome in fact, can be cheap if the cost of the coffee in Rome is compared with that of a coffee in Berlin or expensive if compared with the cost of coffee in Naples. So one needs to know whether the standard of comparison is the cost of coffee in Berlin or Naples. In the case of absolute adjectives, their interpretation is context independent as they map to a scale with a minimal or maximal element which represents the standard of comparison. If something is *dry* (144b) for instance, it possesses the property of absolute dryness whose value is independent of the context.
Context-insensitive adjectives can be further classified in *minimum standard* and *maximum standard* adjectives depending on whether they express the presence of a minimal degree of the property they describe (145a) or its complete absence (145b).

(145) a. *The gold is impure/The table is wet/The door is open*

b. *The platinum is pure/The floor is dry/The door is closed*

Context-sensitive adjectives can be further subdivided into two classes *marked* and *unmarked* depending on whether they can be preceded by an NP expressing the value of a measure or not.

(146) a. *The train is 3 m long/The car is 150 km/h fast* unmarked-form

b. **The train is 1 m short/*The car is 40 km/h slow* marked-form

The notion of markedness corresponds to that of polarity (Bierwisch, 1989), which generally is tested with questions. The unmarked form (e.g. *long*) of the adjective corresponds to a positive polarity item and the marked form (e.g. *short*) to a negative polarity item.

(147) a. *The train is 3 meter long/How long is the train?* polarity:+

b. **The train is 1 meter short/*How short is the train?* polarity:-

In sum, adjectives can be gradable or non gradable. Gradable adjectives can be further subdivided into context-sensitive (or relative) and context-insensitive (or absolute) adjectives. Context-sensitive adjectives can be either marked or unmarked; and context-insensitive adjectives can be either minimum or maximum standard. These categories are summarised in the following table.

gradable	absolute	minimum standard	*wet*
		maximum standard	*dry*
	relative	unmarked	*tall*
		marked	*small*
non gradable			*rectangular*

To classify the sample adjectives with respect to each of these dimensions, the following criteria were used.

Gradability. To decide whether an adjective was gradable or not, I relied for the most part on data discussed in (Kennedy and McNally, 1999; Kennedy, 2004; Kennedy and McNally, 2005; Kennedy, 2005). For these adjectives that were not discussed by Kennedy but were in the sample, I consulted the information found in Wiktionary[32], the on-line dictionary and thesaurus developed by Wikipedia and on this basis, decided whether or not the adjective was gradable.

Context-sensitivity. To distinguish between context-sensitive and context-insensitive gradable adjectives, I relied on the following two syntactic criteria proposed by (Kennedy and McNally, 1999).

Modifiability by *very* Context-sensitive adjectives can be modified by *very*, whereas context-insensitive ones cannot.

(148) a. *The baby is very tall/short/fast*

b. *The glass is very expensive/clean/dirty*

c. **The baby is very awake/asleep*

d. **The glass is very full/empty*

Note that *very* can modify adjectives with open scales (e.g. *big/small*) as well as adjectives which represent the open part of a scale (e.g. *wet* but not *dry*).

(149) *The dishcloth is very wet/clean/dirty/*dry*

For-PP constructions Context-sensitive adjectives are felicitous in for-PP constructions. In fact, such PPs affect the computation of the standard of comparison. In contrast, context-insensitive adjectives do not.

[32] http://www.wiktionary.org/

(150) a. *The baby is tall/short/fast for a two year old*
b. **The baby is awake/asleep for a two year old*
c. **The glass is full/empty for a wine glass*

Markedness. To distinguish between marked and unmarked adjectives, I tested whether for the given adjective the construction `[NP`$_{meas}$ `Adj N]` is felicitous or not.

7.2.4 Syntactic Properties

As syntactic discrimination criteria, I use several meaning preserving alternations of adjectives, namely: (i) predicative/attributive constructions, (ii) *for/as* constructions and (iii) adjectival constructions with clausal complement (SC) such as it-extraposition, and *easy* constructions.

Attributive/Predicative Constructions

English adjectives can be subdivided into adjectives which can only be used predicatively (such as *afloat*), adjectives which can only be used attributively (such as *mechanical* in *mechanical engineer*) and adjectives which can be used in both constructions such as *fast*.

- Attributive Construction
  `This is A N`

  (151) *John is a mechanical engineer.*

- Predicative Construction
  `N is A`

  (152) *The boat is afloat.*

- Predicative/Attributive Alternation
  `This is A N` $\leftrightarrow$ `This N is A`

  (153) *This is a fast car.* $\leftrightarrow$ *This car is fast.*

For/As Constructions

In English, the base attributive form of adjective-noun compounds, i.e. [X is A N], can alternate with *for*-constructions, or with *as*-constructions. As Siegel (1976) notes, a subset of gradable adjectives (such as *big*, *quick*, *expensive*) can alternate with *for*-constructions, but others cannot (154). Still another subset of gradable adjectives (such as *good, cruel, stupid*) can alternate with *as*-constructions but not with *for*-constructions (155).

- For-Construction
  `This is A N` $\leftrightarrow$ `This is A for an N`

(154) a. Bixi is a big mouse. $\leftrightarrow$ *Bixi is big for a mouse.*

b. John is a good cook. $\not\leftrightarrow$ *John is good for a cook.*

- As-Construction
  `This is A N` $\leftrightarrow$ `This is A as an N`

(155) a. John is a good cook. $\leftrightarrow$ *John is good as a cook.*

b. Bixi is a big mouse. $\not\leftrightarrow$ *Bixi is big as a mouse.*

Adjectival Constructions with Sentential/Clausal Complement

Adjectives can be subdivided into different groups depending on their behaviour with respect to complementation:

- Adjectives with Prepositional Complement

(156) *John is fond of his nephew*

- Adjectives with Clausal toVP Complement

(157) *John is loath to admit a mistake*

- Adjectives with Sentential SC Complement

(158) *It is possible that John wins the race*

Following Vendler (1968) and Silva and Thompson (1977), adjectives with sentential and clausal complement can be further subdivided into *object embedding* and *subject embedding* adjectives.

*(159) a. John is eager to come./*To come is eager* ObjE

b. John was brave to come./To come was brave of John SubjE

Adjectives with subject embedding always allow it-extraposition

- It-Extraposition
  `It is A SC` $\leftrightarrow$ `SC is A`

(160) a. John was brave to come. $\leftrightarrow$ *It was brave of John to come*

*b. John is eager to come./*It was eager to come*

Arnold (1989) points out that subject embedding adjectives can be subdivided into five subgroups depending on the preposition which introduces the nominal complement (i.e. the modified noun) thus originating the following paraphrastic pairs:

- S-only, no prepositional complement:
  `It is A SC` $\leftrightarrow$ `SC is A` SubjE(S-only)

(161) *It is possible that they left* ↔ *That they left is possible*

- S-toNP:
  `N is A SC ↔ SC is A to N` SubjE(to)

(162) *Sam was clear that they left* ↔ *That they left was clear to Sam*

- S-ofNP:
  `N is A SC ↔ SC is A of N` SubjE(of)

(163) *John is stupid to take this job* ↔ *To take this job is stupid of John*

- S-forNP:
  `N is A SC ↔ SC is A for N` SubjE(for)

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

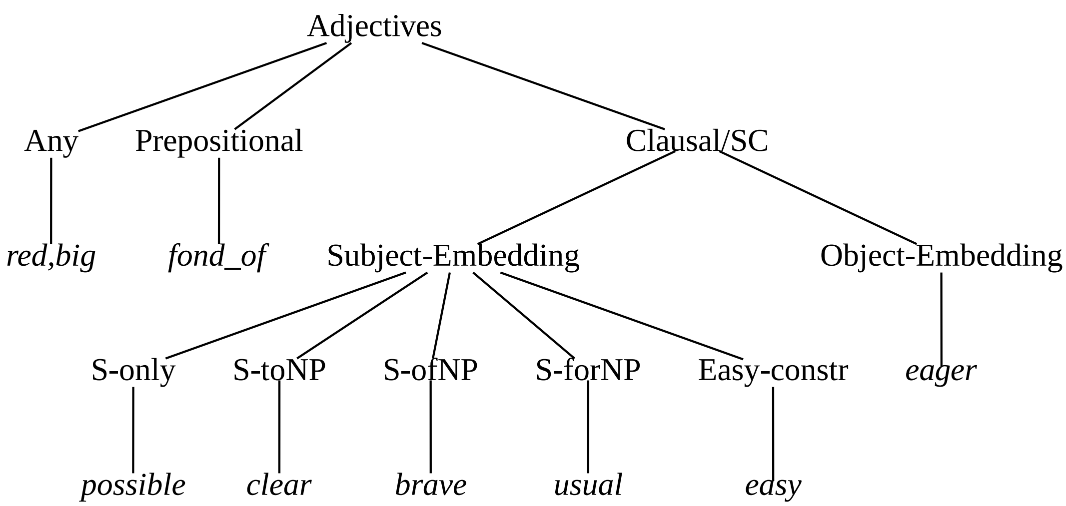


Figure 7.4: Adjectives and Complementation

A subset of adjectives which allow subject embedding can participate in paraphrastic patterns called *easy*-constructions (Chomsky, 1964; Flickinger and Nerbonne, 1992), i.e. constructions in which the modified noun appears as a non-subject complement of the clausal verb.

- Easy-Construction I
  `N is A SC ↔ It is A SC` SubjE(Easy)

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

- Easy-Construction II
  `N is A for-PP SC ↔ It is A for-PP SC` SubjE(Easy)

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

Figure 7.4 summarises the possible complementation of English adjectives.

The assignment of a set of alternations to adjectives was mainly made by relying on the results of other classifications. When the adjective had not been previously classified, an appropriate sentence was submitted as a query for Google. Depending on the number of documents found in the Web containing the given syntactic construction, it was judged felicitous or not. Then, the corresponding variant was constructed and two human raters have judged whether they are synonymous or not.
For example, the adjective *stupid* can participate in both the *of* and *for* constructions but the frequency of the *of*-variant is higher than that of the *for*-variant. Furthermore, only the *of*-variant is paraphrastic with the base form. Hence the adjective *stupid* was classified as accepting the *of*-alternation.

Adjective	Query	Documents Found
stupid	'It was stupid of him to'	722
	'It was stupid for him to'	222

Figure 7.5: Google Results for *of/for*-Constructions

7.2.5 Derivational Morphology

Finally, as derivational morphology selectional criteria, I take into account different paraphrastic constructions in which an adjective is substituted with a morphologically related verb or noun.
For adjective-verb substitutions, I distinguish between cases in which the adjective describes the same event described by the verb (V_a) and cases in which the adjective describes the result of the action described by the verb (V_{res}).

V_a There exists a verb that is semantically related to the adjective.

(167) *John is asleep* $\leftrightarrow$ *John sleeps*

V_{res} The adjective describes the result of the action described by the semantically related verb.

(168) *John has opened the door* $\rightarrow$ *The door is open*

For nominalisations I further distinguish whether the morphologically related noun is an event noun (N_e) or a non-event noun (N_a) or a noun denoting a category N_{cat} or whether the substitution involves a prepositional phrase N_{rel}.

N_e The adjective is semantically related to a noun denoting an event.

(169) *John is asleep* $\leftrightarrow$ *The sleep of John*

N$_a$ The adjective is semantically related to a non-event noun.

(170) *The student is polite* $\leftrightarrow$ *The politeness of the student*

N$_{rel}$ The adjective is substituted with a PP containing the morphologically related noun.

(171) *This is a gastronomical dictionary* $\leftrightarrow$ *This is a dictionary about gastronomy*

N$_{cat}$ The adjective is substituted with a morphologically related noun denoting a category of individuals.

(172) *This animal is carnivorous* $\leftrightarrow$ *This animal is a carnivore*

Information about the morphological properties of adjectival items was taken from WordNet. For each adjective, the morphologically related nouns and verbs were extracted, then sentences were constructed with follow the syntactic patterns described above and judged by two human raters for being paraphrastic or not.

(173) a. N is A $\leftrightarrow$ *N Va*

b. John is asleep $\leftrightarrow$ John sleeps

c. the table is red $\not\leftrightarrow$ the table reddens

In case of nonagreement the pattern was rejected for the given adjective.
For N$_{rel}$, I further relay on the work of Levi (1978) on nominal compounds.

7.3 Adjectival Classes

By applying the criteria described in section 7.1 to the sample set of 500 adjectives, I have obtained the classification shown in figure 7.6, which defines 42 adjectival classes. I use the following notational convention for the symbols used in the figures:

Adjectives

MT	Abs	Trans	Grad	very	Anto	Stab	P/A	for/as	SCcomp	NPMeas	MorphoD	Class Name	
I	A	T	G	-	M	S	PA			-	N_a	*red1*	**Ipa1**
I	A	T	G	-	B	D	P			-	V_a, N_e	*asleep1*	**Ip1**
I	A	T	G	-	B	D	PA			-	N_a	*dry1*	**Ipa2**
I	A	T	G	-	B	S	PA			-	N_a	*straight1*	**Ipa3**
I	A	T	G	+	B	S	PA			-	N_a	*rough1*	**Ipa4**
I	A	T	G	+	B	D	PA			-	N_a	*ill1*	**Ipa5**
I	A	T	NG	-	M	D	P			-	V_a, N_e	*afloat2*	**Ip2**
I	A	T	NG	-	M	S	PA			-	N_{cat}	*carnivorous1*	**Ipa6**
I	A	T	NG	-	M	S	PA			-	N_a	*rectangular1*	**Ipa7**
I	A	T	NG	-	M	S	PA			-	N_{rel}	*artisanal1*	**Ipa8**
I	A	T	NG	-	M	S	A			-	N_{rel}	*present1*	**Ia1**
I	A	T	NG	-	B	S	PA			-	N_a	*finite1*	**Ipa9**
I	A	T	NG	-	B	S	PA		SubjE(S-only)	-	N_a	*real1*	**Ipa10**
I	A	T	NG	-	B	D	PA			-	V_{res}	*closed1*	**Ipa11**
I	A	T	NG	-	B	D	PA			-	N_a	*absent1*	**Ipa12**
S	NA	T	NG	-	M	S	A			-	N_{rel}	*gastronomical1*	**Sa1**
S	NA	T	NG	-	C	S	PA	for		-	N_a	*enormous1*	**Spa1**
S	NA	T	G	+	C	S	PA	for		+	N_a	*big1*	**Spa2**
S	NA	T	G	+	C	S	PA	for		-	N_a	*small1*	**Spa3**
S	NA	T	G	+	C	D	PA	for		+	N_a	*fast1*	**Spa4**
S	NA	T	G	+	C	D	PA	for		-	N_a	*slow1*	**Spa5**
S	NA	NT	G	+	B	D	PA	as	SubjE(Easy)	-	N_a	*useful1*	**Spa6**
S	NA	NT	G	+	B	D	PA	as	SubjE(for)	-	N_a	*certain2*	**Spa7**
S	NA	NT	G	+	B	D	PA	as	SubjE(S-only)	-	N_a	*necessary1*	**Spa8**
S	NA	NT	G	+	B	D	PA	as	SubjE(to)	-	N_a	*clear1*	**Spa9**
S	NA	NT	G	+	B	D	PA	as	ObjE	-	N_a	*eager1*	**Spa10**
S	NA	NT	G	+	C	S	PA	as	SubjE(S-only)	-	N_a	*intelligent1*	**Spa11**
S	NA	NT	G	+	C	D	PA	as	SubjE(Easy)	-	N_a	*easy1*	**Spa12**
S	NA	NT	G	+	C	D	PA	as	SubjE(of)	-	N_a	*cruel1*	**Spa13**
S	NA	NT	G	+	C	D	PA	as	SubjE(for)	-	N_a	*sad1*	**Spa14**
S	NA	NT	G	+	C	D	P	as	PP, ObjE	-	N_a	*averse1*	**Sp1**
S	NA	NT	NG	-	C	D	A			-		*main1*	**Sa2**
S	NA	NT	NG	-	C	S	A			-		*utter1*	**Sa3**
S	NA	NT	NG	-	M	D	A			-		*second1*	**Sa4**
PR	A	T	NG	-	B	S	PA		SubjE(S-only)	-	N_a	*fictitious1*	**PRpa1**
PR	NA	NT	NG	-	B	S	PA			-	N_a, N_{cat}	*fake1*	**PRpa2**
PR	NA	NT	NG	-	B	S	A			-	V_a	*pretended1*	**PRa1**
PR	NA	NT	NG	-	B	D	PA	as	SubjE(For)	-	V_{res}	*foreseen1*	**PRpa3**
PR	NA	NT	NG	-	B	D	A		ObjE	-	N_a	*potential1*	**PRa2**
PR	NA	NT	NG	-	M	S	A			-	N_{rel}	*former2*	**PRa3**
PR	NA	NT	G	+	B	D	PA	as	SubjE(S-only)	-	N_a	*probable2*	**PRpa4**
NS	NA	NT	NG	-	B	D	A		SubjE(of)	-	V_{res}	*alleged2*	**PlNS**

Figure 7.6: Classes of English Adjectives

Model Theoretic Properties:	
MT	the model theoretic group can be instantiated with I(ntersective), S(ubsective), PR(ivative), NS(plain nonsubsective),
Abs	absoluteness can be instantiated with A(bsolute) or NA(nonabsolute),
Trans	transparency can be instantiated with T(ransparent) or NT(nontransparent),
Lexical Semantic Properties:	
Grad	gradability can be instantiated with G(radable) or NG(nongradable)
Very	*very*-modification can be instantiated with +(can be modified by *very*) or -(cannot be modified by *very*)
Anto	antonymy can be instantiated with B(inary), C(ontrary) or M(ulti-opposition)
Stab	stability can be instantiated with S(table) or D(ynamic)
Syntactic Properties:	
P/A	adjective usage can be instantiated with PA(both predicative and attributive usage), P(only predicative usage), A(only attributive usage)
for/as	*for/as*-construction can be instantiated with 'for'(the adjective allows *for*-construction), 'as'(the adjective allows *as*-construction) or is empty (the adjective does not allow anyone of these constructions)
SCcomp	constructions with sentential complement can be instantiated with ObjE (the adjective allows object embedding) or SubjE (the adjective allows subject embedding) of type S-only (without prepositional complement) or to/of/for depending on the preposition in the prepositional complement
NPMeas	the adjective (+) can be preceded by an NP denoting a measure value or not (-)
Morphoderivational Properties:	
MorphoD	can be instantiated with N_a, N_e, N_{cat}, N_{rel}, V_a,V$_res$ which have the meaning described in section 7.2.5

For simplicity of description, I start by the four main semantic classes defined by Kamp and Partee, i.e. *intersective*, *subsective*, *privative* and *plain nonsubsective* and I show how by taking into account the full set of properties of adjectives previously discussed, namely (i) model theoretic, (ii) lexical semantic, (iii) syntactic and (iv) morphoderivational properties, each main class can be further subdivided into subclasses with homogeneous inferential behaviour.

7.3.1 Intersective Adjectives

Traditionally, adjectives classified as *intersective* are those adjectives that form $[\mathrm{A\ N}]_{AP}$ phrases from which it is possible to infer both the property expressed by the noun N and the property expressed by the adjective A. This class includes adjectives such as *rectangular, real, open, dry, wet*. In the present classification, I also include as members of the intersective class adjectives such as *afloat, asleep* which syntactically can occur only in predicative position and adjectives such as *present* (in temporal sense) which syntactically can only be used attributively as they share the model theoretic properties of intersective adjectives, namely *subsectivity*, *absoluteness* and *transparency*.

The application of the test criteria proposed in section 7.2 to this set of adjectives shows that they do not form a homogeneous class.

Intersective Adjectives											
MT	Abs	Trans	Grad	very	Anto	Stab	P/A	SCcomp	MorphoD	Class	
I	A	T	G	-	M	S	PA		N_a	*red1*	**Ipa1**
I	A	T	G	-	B	D	P		V_a,N_e	*asleep1*	**Ip1**
I	A	T	G	-	B	D	PA		N_a	*dry1*	**Ipa2**
I	A	T	G	-	B	S	PA		N_a	*straight1*	**Ipa3**
I	A	T	G	+	B	S	PA		N_a	*rough1*	**Ipa4**
I	A	T	G	+	B	D	PA		N_a	*ill1*	**Ipa5**
I	A	T	NG	-	M	D	P		V_a,N_e	*afloat2*	**Ip2**
I	A	T	NG	-	M	S	PA		N_{cat}	*carnivorous1*	**Ipa6**
I	A	T	NG	-	M	S	PA		N_a	*rectangular1*	**Ipa7**
I	A	T	NG	-	M	S	PA		N_{rel}	*artisanal1*	**Ipa8**
I	A	T	NG	-	M	S	A		N_{rel}	*present1*	**Ia1**
I	A	T	NG	-	B	S	PA		N_a	*finite1*	**Ipa9**
I	A	T	NG	-	B	S	PA	SubjE(S-only)	N_a	*real1*	**Ipa10**
I	A	T	NG	-	B	D	PA		V_{res}	*closed1*	**Ipa11**
I	A	T	NG	-	B	D	PA		N_a	*absent1*	**Ipa12**

Figure 7.7: Subclasses of Intersective Adjectives

As Figure 7.7 shows, lexical semantics, syntactic and morphoderivational selectional criteria individuate 15 subclasses of intersectives. For example, by considering their behaviour with respect to gradability, intersective adjectives can be subdivided into two subclasses, the class of gradable (e.g. *red, dry, ill*) and the class of nongradable (e.g. *afloat, finite, closed*) intersective adjectives.

(174) a. *The soil is wetter than usual.*

b. **This is more rectangular than usual*

The set of gradable intersectives can be further subdivided in adjectives which can be modified by *very* and those which cannot.

(175) a. *John is very ill*

b. **This door is very red*

Antonymy defines other subclasses. The adjectives *red* and *dry* for example are both gradable and cannot be modified by very, but whereas *dry* enters in a binary antonymic relation with its antonym, *red* enters in multiple oppositions.

(176) a. *The soil is not dry* $\leftrightarrow$ *The soil is wet*

b. *This door is not red* $\not\rightarrow$ *This door is blue*

By further considering stability/dynamicity criteria, intersective adjectives can be subdivided into the set of stable (e.g. *carnivorous, rectangular, finite*) and into the set of dynamic (e.g. *asleep, dry*) intersectives.

(177) a. *John is being ill./The door is being open.*

b. **The table is being rectangular.*

A still finer grained classification of intersective adjectives can be obtained by taking into consideration their derivational morphology (178).

(178) a. *John is asleep.* $\leftrightarrow$ *John sleeps.* Adj $\rightarrow$ V_a

b. *John has closed the door.* $\rightarrow$ *The door is closed.* Adj $\rightarrow$ V_{res}

c. *John is ill.* $\leftrightarrow$ *John's illness.* Adj $\rightarrow$ N_a

d. *This animal is carnivorous.* $\leftrightarrow$ *This animal is a carnivore.* Adj $\rightarrow$ N_{cat}

The 15 subclasses of intersective adjectives are described below.

Class Ipa1 Members of this class are adjectives such as *blue1, black1, green1, grey1, orange1, red1, white1, yellow1*. These adjectives describe colour properties. They are gradable but cannot be modified by *very*. They denote static properties and enter in multi-opposition relations with their antonyms.

Class Ipa2 Members of this class are adjectives such as *dry1, empty1, full1, lighted1*. These gradable adjectives correspond to the context insensitive adjectives described in Kennedy (2005). They cannot be modified by *very* and in fact represent the extreme values, i.e. punctual values, of a closed or partially closed scale. Further, the adjectives in this class denote dynamic properties and enter in binary relations with their antonyms.

Class Ipa3 Members of this class are adjectives such as *anterior1, backward1, convergent1, divergent1, forward1, posterior1, straight2, unlighted1, unbreakable1*. These gradable adjectives denote context insensitive properties. They cannot be modified by *very* and represent the extreme values of a closed or partially closed scale. But, contrary to the adjectives in the Ipa2 class, they denote stable properties.

Class Ipa4 Members of this class are adjectives such as *adhesive1, alcoholic11, curly1, flat1, fluffy1, fragile1, fragrant1, greasy1, impure1, musical2, nonadhesive1, nonalcoholic1, nonfat1, pointless1, pure1, raw3, rough1, sharp2, smelly1, smooth1, sour2, spicy1, spiny1, sweet1, tasteful1, tasteless1, uneven1, unmusical1, viscid1*. These adjectives denote stable, context insensitive properties, can be modified by *very* and enter in binary antonymic relations with their antonyms.

Class Ipa5 Members of this class are adjectives such as *bright1, clean1, clear11, cloudy3, dirty1, drunk1, dull2, dusty1, garrulous1, good-natured1, healthy1, hungry1, ill1, irritable1, noisy1, nonslippery1, peaceful1, placid1, quiet2, ripe1, sick1, satiate1, shaky1, slippery1, sober1, stable1, stormy1, sunny1, taciturn1, tame3, testy1, thirsty2, tidy1, unhealthy1, unpeaceful1, unripe1, well1, wet1, wild2*.
These adjectives similarly to the adjectives in the precedent class denote context insensitive properties. They can be modified by *very* and enter in binary antonymic relations with their antonyms. But contrary to the members of the Ipa4 class, they denote dynamic properties.

Class Ipa6 Members of this class are adjectives such as *androgynous1, carnivorous1, female1, gaseous1, herbivorous1, insectivorous1, liquid1, male1, omnivorous1 solid1*. These adjectives are nongradable, denote stable properties and enter in multiple opposition relations with their antonyms. Their principal characteristic is that they can be nominalised by direct substitution with the morphologically related noun. These adjectives define species and state of matter.

Class Ipa7 Members of this class are adjectives such as *oval1, quadrate1, rectangular1, round1, triangular1*. Such adjectives are denote shapes. They are stable and enter in multiple opposition relations with their antonyms.

Class Ipa8 Members of this class are adjectives such as *artisanal1, bacterial1, biennial1, biologic1, bisyllabic1, constitutional1, cultural1, economic1, immediate4, metallic1, monthly1, philatelic1, weekly, wooden1, yearly1*. These adjective corresponds to the relational adjectives described in Levi (1978), but they can also be used predicatively.

Class Ipa9 Members of this class are adjectives such as *blind1, blond1, brunet1, deaf1, edible1, even1, finite1, fractional1, infinite1, mute3, odd1, partial1, permanent1, spineless1, temporary1, unique3, vacuous3, various1*. This class includes adjectives which represent stage-level predicates. They are nongradable and enter in binary opposition relations with their antonyms.

Class Ipa10 Members of this class are adjectives such as *actual2, authentic2, genuine1, natural2, real1*. This class includes adjectives which are known in the literature as tautological as they express the obvious property of existence or realness of all entities and eventualities to which they are applied: $\forall x : Adj(x) = true$[33]. Syntactically, they can be used both predicatively and attributively and are not gradable. Semantically, they enter in binary opposition relations with their antonyms and describe static properties. It is worth noting that the antonyms of the adjectives in this class are all privative.

[33]This is true of everything which is not modified by a privative adjective, a fictitious hero is not real nor existent.

Class Ipa11 Members of this class are adjectives such as *accompanied1, barred1, bent1, broken1, closed1, clothed1, cooked1, crooked1, damaged1, dead1, dressed1, fastened1, hand-made1, locked1, machine-made1, married1, open1, scared1, shut1, unbroken1, unclothed1, undamaged1, unfastened1, unmarried1*. This gradable intersective adjectives have the property of entering in binary oppositions relations with their antonyms and represent dynamic properties. They are morphologically related to verbs.

Class Ipa12 Members of this class are adjectives such as *absent1, dynamic3, motionless1, omnipresent1, present2, static1, ubiquitous1*. The adjectives in this class denote dynamic properties and are nongradable.

Class Ip1 Members of the class are adjectives such as *ablaze2, afire1, aflame2, aglow1, afraid1, alive1, asleep1, awake1*. These adjectives can only be used predicatively. They are morphologically related to verbs and event nouns, are gradable and represent dynamic properties.

Class Ip2 Members of the class are adjectives such as *adrift2, afloat2, aground1*. They can only be used predicatively. Such adjectives are morphologically related to verbs, are not gradable and represent dynamic properties.

Class Ia1 Members of the class are adjectives such as $present_1$. These adjectives denote the existence of an individual at the present time and thus can be included in the tautological adjective. Syntactically, this class includes adjectives which have only attributive usage. Semantically, such adjectives enter in multiple opposition relations with their antonyms, are transparent and describe dynamic properties.

7.3.2 Subsective Adjectives

This class traditionally includes adjectives such as *big, good, mechanical, recent* that form $[A\ N]_{AP}$ phrases from which it is only possible to infer the property expressed by the noun but not the one expressed by the adjective. As shown in Figure 7.8, this general class can be subdivided in 19 more specific subclasses.
By applying model theoretic selectional criteria to the adjectives in this class, they can be subdivided into two big groups, the subsective adjectives which are transparent and those which are not. The first group, the subsective transparent adjectives, includes adjectives such as *big, small* which denote dimensional properties and relational adjectives such as *wooden, gastronomical, solar*.
The second group, the subsective nontransparent adjectives, includes members such as *easy, cruel, main, utter* which denote nondimensional properties.
The set of transparent adjectives denoting dimensional properties can be further subdivided into five subclasses by considering gradability, staticity/dynamicity and syntactic properties:

(179) a. *A mouse is bigger than an ant/*This building is more enormous than that one.*

Subsective Adjectives												
MT	Abs	Trans	Grad	very	Anto	Stab	P/A	for/as	SCcomp	NPMeas	MorphoD	Class
S	NA	T	NG	-	M	S	A			-	N_{rel}	*gastronomical1* **Sa1**
S	NA	T	NG	-	C	S	PA	for		-	N_a	*enormous1* **Spa1**
S	NA	T	G	+	C	S	PA	for		+	N_a	*big1* **Spa2**
S	NA	T	G	+	C	S	PA	for		-	N_a	*small1* **Spa3**
S	NA	T	G	+	C	D	PA	for		+	N_a	*fast1* **Spa4**
S	NA	T	G	+	C	D	PA	for		-	N_a	*slow1* **Spa5**
S	NA	NT	G	+	B	D	PA	as	SubjE(Easy)	-	N_a	*useful1* **Spa6**
S	NA	NT	G	+	B	D	PA	as	SubjE(for)	-	N_a	*certain2* **Spa7**
S	NA	NT	G	+	B	D	PA	as	SubjE(S-only)	-	N_a	*necessary1* **Spa8**
S	NA	NT	G	+	B	D	PA	as	SubjE(to)	-	N_a	*clear1* **Spa9**
S	NA	NT	G	+	B	D	PA	as	ObjE	-	N_a	*eager1* **Spa10**
S	NA	NT	G	+	C	S	PA	as	SubjE(S-only)	-	N_a	*intelligent1* **Spa11**
S	NA	NT	G	+	C	D	PA	as	SubjE(Easy)	-	N_a	*easy1* **Spa12**
S	NA	NT	G	+	C	D	PA	as	SubjE(of)	-	N_a	*cruel1* **Spa13**
S	NA	NT	G	+	C	D	PA	as	SubjE(for)	-	N_a	*sad1* **Spa14**
S	NA	NT	G	+	C	D	P	as	PP, ObjE	-	N_a	*averse1* **Sp1**
S	NA	NT	NG	-	C	D	A			-		*main1* **Sa2**
S	NA	NT	NG	-	C	S	A			-		*utter1* **Sa3**
S	NA	NT	NG	-	M	D	A			-		*second1* **Sa4**

Figure 7.8: Subclasses of Subsective Adjectives

b. *The train is 2 m long/*John is 1.50 m small*

c. *This is a mechanical engineer/*This engineer is mechanical*

d. *This is the present/recent president/*This president is present/recent*

e. *John is being fast/*John is being fat*

The set of nondimensional adjectives can be subdivided into two big subgroups by distinguishing between gradable (e.g. *cruel, necessary*) and nongradable adjectives (e.g. *main, utter*).
The group of gradable nontransparent subsectives can be further subdivided into 10 subclasses by considering antonymy, stability/dynamicity criteria and specifically the type of SC-complement these adjectives allow:

(180) a. *John is eager to read the book/*To read the book is eager* ObjE

b. *The king was cruel to condemn the boy* ↔ *To condemn the boy was cruel of the king* SubjE(of)

c. *John is easy to talk to* ↔ *It is easy to talk to John* SubjE(Easy)

d. *John was sad to leave* ↔ *To leave was sad for John* SubjE(for)

e. *John was clear that the problem couldn't be solved* ↔ *It was clear to John that the problem couldn't be solved* SubjE(to)

A further distinction between the nongradable subsective adjectives can be made by applying lexical semantic criteria, e.g. antonymy and stability/dynamicity.
Thus, adjectives such as *main, utter* enter in contrary opposition with their antonyms whereas adjectives such as *second* enter in multiple oppositions.

Class Spa1 Members of this class are adjectives such as *atomic3, average1, enormous1, giant1, immense1, intermediate2, medium1, oceanic2, superb2, tepid1*.
Such adjectives are dimensional but nongradable and describe stable properties.

Class Spa2 Members of this class are adjectives such as *aged1, big1, deep3, fat1, heavy1, high1, huge1, large1, long1, narrow1, old1, tall1, thick1, wide1*.
These adjectives which traditionally are classified as scalar, describe dimensional comparable properties with open scales. This set includes the marked forms (Kennedy, 2005) of the adjectives. They can be modified by *very* and by NP providing measure values. Further, the adjectives in this class denote stable properties.

Class Spa3 Members of this class are adjectives such as *colossal1, gigantic1, great1, hard3, light1, little1, low1, miniature1, minuscule3, monumental1, shallow1, short3, short1, small1, soft1, thin2, vast1, young1*.
These adjectives which traditionally are classified as scalar, describe dimensional

comparable properties with open scales. This set includes the unmarked forms (Kennedy, 2005) of the adjective. They can be modified by *very*, by NP providing measure values and denote stable properties.

Class Spa4 Members of this class are adjectives such as *expensive1, fast1, hot1, late2, loud1, quick1, rich1, strong1, warm1, wealthy1*.
These adjectives which traditionally are classified as scalar, describe dimensional comparable properties with open scales. This set includes the marked forms (Kennedy, 2005) of the adjective. They can be modified by *very*, and by NP providing measure values but contrary to the adjectives in the Spa2 class, they denote dynamic properties.

Class Spa5 Members of this class are adjectives such as *cheap1, cold1, cool1, icy2, gelid1, poor2, shrill1, slow1, soft3, swift1, weak1*.
These adjectives which traditionally are classified as scalar, describe dimensional comparable properties with open scales. This set includes the unmarked forms (Kennedy, 2005) of the adjective. They can be modified by *very*, and by NP providing measure values but contrary to the adjectives in the Spa3 class they denote dynamic properties.

Class Spa6 Members of this class are adjectives such as *dangerous1, pleasant1, profitable1, safe1, unpleasant1, unprofitable1, useful1, useless1*.
This class includes gradable adjectives which do not describe dimensional properties. These adjectives are gradable and enter in binary antonymic relations with their antonyms. Further, they can participate in subject embedding (easy) constructions.

Class Spa7 Members of this class are adjectives such as *certain2, frequent1, habitual1, infrequent1, sure1, rare2, unusual1, usual1*.
This class includes gradable adjectives which do not describe dimensional properties. These adjectives are gradable and enter in binary antonymic relations with their antonyms. They can participate in subject embedding constructions (preposition *for*).

Class Spa8 Members of this class are adjectives such as *acknowledged1, appropriate1, arbitrary1, damaging1, decisive1, famous1, known1, inappropriate1, indecisive2, insufficient1, necessary1, scarce1, sufficient1, true1, unacknowledged1, uncontroversial1, unknown1, unnecessary1*.
This class includes gradable adjectives which do not describe dimensional properties. These adjectives are gradable and enter in binary antonymic relations with their antonyms. They can participate in subject embedding constructions (S-only).

Class Spa9 Members of this class are adjectives such as *clear1, evident1, new1, novel2, obscure1, old2, obvious1, recent1, unclear1*.
This class includes gradable adjectives which do not describe dimensional properties. These adjectives are gradable and enter in binary antonymic relations

with their antonyms. They can participate in subject embedding constructions (preposition *to*).

Class Spa10 Members of this class are adjectives such as *angry1, anxious2, desperate1, disinclined1, eager1, experienced1, impatient1, inexperienced1, jealous2, patient1, ready1, unangry1, uneager1, unready1, unwilling1, willing1*.
They describe gradable adjectives with totally open scales. Such adjectives participate in binary opposition relations with their antonyms and in object embedding constructions.

Class Spa11 Members of this class are adjectives like *atypical1, awkward2, beautiful1, clumsy1, graceful1, handsome1, hopeful1, hopeless1, intelligent1, pretty1, serious1, skilled1, skillful1, unskilled1, typical1, ugly1, unexpected1*.
This class includes gradable adjectives which do not describe dimensional properties. These adjectives are gradable and enter in binary antonymic relations with their antonyms. They can participate in subject embedding constructions (S-only).

Class Spa12 Members of the class are adjectives such as *amusing2, bad1, comfortable1, complex1, difficult1, easy1, excellent1, exciting1, funny1, good1, hard1, important1, interesting1, nice1, perfect1, satisfying1, simple1, tough2, tricky2, unamusing1, uncomfortable1, unexciting2, unimportant1, uninteresting1, unsatisfying1*.
This class corresponds to the class of the *easy* adjectives (Flickinger and Nerbonne, 1992). This class is ontologically quite homogeneous and includes properties denoting evaluation of difficulty.

Class Spa13 Members of this class are adjectives like *altruistic1, ambitious1, artful1, brave1, capricious1, civil2, clever3, considerate1, crazy1, criminal3, cruel1, disloyal2, egoistic1, foolish1, friendly1, generous1, humane1, immoral1, impolite1, inconsiderate1, ingenuous1, inhumane1, insane1, insensible1, just1, kind1, loyal3, polite1, rude1, sensible1, stingy1, stupid1, tender1, tough1, unfriendly1, unjust2, unkind1, unreasonable1, wise1*.
This class includes properties which denote an evaluation of an individual. They represent gradable properties and participate in subject embedding constructions (preposition of).

Class Spa14 Members of this class are adjectives such as *fortunate1, happy1, sad1, unhappy1, unfortunate1*.
They describe psychological states. These adjectives are gradable, dynamic and enter in contrary antonymic relations with their antonyms. Further, they participate in subject embedding constructions (preposition for).

Class Sa1 Members of this class are adjectives such as *acid3, animal1, aquatic1, atomic1, chemical1, civil1, criminal0, dental1, electrical1, feminine1, financial1, gastronomical1, linguistic1, lunar1, marginal1, marine1, masculine1, maternal1, mathematical1, mechanical3, medical1, metal1, molecular1, moral1, musical1, na-*

tional1, nuclear2, oceanic1, paternal1, plastic1, polar4, presidential1, provincial1, solar1, stellar2, urban1, viral1.
These adjectives are relational (Levi, 1978) and are derived from nouns. They denote nongradable properties and can only be used attributively.

Class Sa2 Members of this class are adjectives such as *main1, primary3, principal1*.
These adjectives are nongradable and denote dynamic properties. They define a ranking for the property expressed by the modified noun.

Class Sa3 Members of this class are adjectives such as *complete4, consummate3, inveterate1, mere1, perfect2, sound8, utter1, veteran1, whole1*.
These adjectives are nongradable and denote static properties.

Class Sa4 Members of this class are adjectives such as *first1, former1, intermediate1, last2, latter1, second1, third1*.
These adjectives are nongradable and denote dynamic properties. They define a ranking for the property expressed by the modified noun.

Class Sp1 Members of the class are adjectives such as *averse1, aware1, fond4, inclined1, indisposed2, disinclined1, loath2, reluctant1*. They can only be used predicatively. Such adjectives are all morphologically related to non-event nouns. They are gradable and represent dynamic properties.

7.3.3 Privative Adjectives

Adjectives such as *former2, fake1, fictitious1, pseudo* are traditionally classified as privatives as they allow the inference pattern:

$$A\ N \rightarrow \neg\ N$$

But, as shown in figure 7.9, privative adjectives do not constitute an homogeneous class.

Privative Adjectives							
MT	Abs	Trans	Grad	Anto	Stab	P/A	Class
PR	A	T	NG	B	S	PA	*fictitious1* **PRpa1**
PR	NA	NT	NG	B	S	PA	*fake1* **PRpa2**
PR	NA	NT	NG	B	S	A	*pretended1* **PRa1**
PR	NA	NT	NG	B	D	PA	*foreseen1* **PRpa3**
PR	NA	NT	NG	B	D	A	*potential1* **PRa2**
PR	NA	NT	NG	M	S	A	*former2* **PRa3**
PR	NA	NT	G	B	D	PA	*probable2* **PRpa4**

Figure 7.9: Subclasses of Privative Adjectives

The first important distinction can be made by applying model theoretic criteria. As example (181) shows, there are privative adjectives such as *fictitious* which are absolute and transparent and privative adjectives such as *fake* which are nonabsolute and nontransparent.

(181) a. *X is a fictitious hero $\rightarrow$ X is not a hero*

b. *X is a fake gun $\rightarrow$ X is not a gun*

c. *X is a fictitious hero$\rightarrow$ X is not an entity*

d. *X is a fake gun $\not\rightarrow$ X is not an object*

By testing for gradability it is possible to identify other subgroups of privative, namely the gradable privatives (e.g. *fictitious, probable*) and the group of nongradable privatives (e.g. *foreseen, pretended*).

(182) a. *This result is more probable./This result is very probable.*

b. **John is more foreseen as president than Mary*

By considering antonymy, the group of nongradable privatives (e.g. *fake*, *foreseen*) which enter in binary opposition with their antonyms can be distinguished from the group of nongradable privatives (e.g. *former*) which have multiple antonyms.

(183) a. *Victor is a former Catholic $\not\rightarrow$ Victor is a future Catholic*

b. *This fur is not fake $\rightarrow$ This fur is authentic*

A more fine grained distinction can be made by distinguishing between stable and dynamic privative adjectives.

(184) a. *Turkey is being foreseen as the command base for the operation.*

b. **This fur is being fake*

The 7 subclasses of privative adjectives obtained by applying the criteria of section 7.2 are described below.

Class PRpa1 Members of this class are adjectives such as *fabricated1, fabulous2, fanciful2, fictional2, fictitious1, fictitious2, fictive1, imaginary1, invented2, legendary2, mythical1, mythic1, nonexistent1, unreal1*.
These adjectives are absolute and transparent and are used to negate the existence (in the real world) of the individual denoted by the noun they modify.

Class PRpa2 Members of this class are adjectives such as *apparent2, artificial1, assumed1, bogus1, counterfeit1, fake1, false6,9, imitative3, inauthentic1, ostensible2, simulated1, spurious3*.

The adjectives in this class are nontransparent, nongradable and denote stable properties. These adjective are used to negate some specific properties of the individual denoted by the noun they modify but not its existence.

Class PRpa3 Members of this class are adjectives such as *expected1, foreseen1, predicted1*.

The adjectives in this class are nongradable and denote dynamic properties. They denote modal properties.

Class PRpa4 This class describes adjectives like *possible1, controversial1, impossible1, improbable1, likely1, probable2, uncertain1, unlikely1*.

These adjectives are gradable and denote dynamic and modal properties.

Class PRa1 Members of this class are adjectives such as *pretended1, pseudo1, would-be1, seeming1*.

These adjectives are nongradable and denote stable properties. They denote the intension of an individual to fake something.

Class PRa2 This class describes adjectives like *eventual1, potential1*.

These adjectives are nongradable and denote dynamic and modal properties.

Class PRa3 Members of this class are adjectives such as *early3, former2,3, future1, incoming1, last1, last2, late3, outgoing1, past1, preceding1, previous1, recent2, succeeding1*.

Such adjectives are used to constrain the time of existence of the property expressed by the noun they modify. It is worth noting, that these adjectives can only be used to modify nouns which express stage-level properties (e.g. *former president, member, student*/**former man, violin, table, glass*).

7.3.4 Plain Nonsubsective Adjectives

Plain nonsubsective adjectives include members such as *alleged, putative*. The main semantic feature of such adjectives is that the $[\mathrm{A\ N}]_{AP}$ combination does not allow to infer neither the property expressed by the adjective nor the property expressed by the noun. Plain nonsubsective adjectives are in fact nonabsolute, nontransparent and nonprivative.

Class PlNS This class describes adjectives like *alleged2, purported1, putative1, reputed1, supposed3*.

The adjectives in this class can only be used attributively, represent dynamic properties, are non gradable and cannot be modified by *very*, they also cannot be nominalised and are morphologically related with verbs belonging to Levin's *Say*-verb class.

Plain Nonsubsective Adjectives							
MT	Abs	Trans	Grad	Anto	Stab	MorphoD	Class
NS	NA	NT	NG	B	D	V_{res}	*alleged2* **PlNS**

Figure 7.10: Subclasses of Intersective Adjectives

7.4 Discussion

The semantic classification presented in this chapter builds on previous classifications by integrating within a single classification scheme, a set of selectional criteria that stems from different levels of linguistic knowledge such as model theoretic semantics, lexical semantics, syntax and morphoderivational knowledge. As a result, this classification defines a set of semantically fine grained classes which, as we shall see in the next chapters, can be used to support the recognition of adjective based textual entailment.

Although the proposed classification is not comprehensive (the coverage is restricted to 500 items and the set of inferential patterns considered does not cover the whole inferential properties of adjectives), it exhibits some interesting properties. In particular, it suggests a systematic relation between ontological and semantic classes. Thus, Figure (7.11) shows the relation between the classification I propose and the taxonomies of adjectives presented in Chapter 6 in particular that of Dixon and Aarts. As can be seen, the distribution of the ontological categories in the semantic classes is quite homogeneous. The intersective adjectives represent non-dimensional physical properties. Subsective adjectives include dimensional physical properties and non-physical properties representing psychological states, evaluations, etc. Privatives and plain nonsubsective adjectives represent non-physical properties, the former corresponding to modal categories.

These considerations provide some evidence that the methodology used to classify adjectives in this thesis namely, the definition of semantic classes by means of an extensive set of criteria, permits establishing a link between taxonomical and linguistic classes and thus, is promising in supporting the creation of general linguistic based ontologies (Bateman, 1990) which may serve as syntax-semantic interfaces to domain oriented knowledge ontologies.

7.5 Summary

In this chapter, I presented a methodology for the classification of adjectives which is based on the approach adopted by Levin (1993) to classify verbs. I have shown how by considering a more extended set of adjectival features, and in particular by integrating syntax and semantic criteria with lexical semantic and morphoderivational ones, and by defining the adjectival classes modulo set of allowed inference patterns, it was possible to build a more fine grained classification, which, compared with previous ones, better captures the different facets of adjectival meaning.

Intersective	
Non-Dimensional Physical Properties	
Shapes	Ipa7 (e.g. *rectangular*)
Taste	Ipa4 (e.g. *spicy*)
Colour	Ipa1 (e.g. *red*)
Physical Properties	Ipa2, Ipa3 (e.g. *dry,wet*)
Corporal Function	Ipa9 (e.g. *blind*)
Corporal Condition	Ip1 (e.g. *asleep*)
Activity	Ip2 (e.g. *afloat*)
State	Ipa11 (e.g. *closed*)
Subsective	
Dimensional Physical Properties	
Size	Spa1, Spa2, Spa3 (e.g. *big, large*)
Speed	Spa4, Spa5 (e.g. *fast, slow*)
...	
Non-Physical Properties	
Human Propensity	
Emotion	Spa10, Sp1 (e.g. *eager, loath*)
Attitude	Spa13 (e.g. *cruel, polite*)
Intellect	Spa11 (e.g. *intelligent*)
Evaluation	Spa12 (e.g. *easy, difficult*)
Necessity	Spa8 (e.g. *necessary*)
...	
Privative	
Non-Physical Properties	
Qualification	
Nonexistence	PRpa1 (e.g. *fictitious*)
Appearance	PRpa2 (e.g. *fake*)
Time	PRa3 (e.g. *early, former*)
Possibility	PRpa4, PR2 (e.g. *foreseen, possible*)
Plain Nonsubsective	
Non-Physical Properties	
Communication	PlNS (e.g. *alleged, putative*)

Figure 7.11: Mapping Adjectival Classes to Adjectival Taxonomies

In the next chapter, I show how adjectival classes can be assigned a first order semantic representation which helps modeling their inferential behaviour.

Assigning a FOL Representation to Adjectival Classes

In the previous chapter, I have shown that the inferential behaviour of adjectives is determined by the interaction of a broad set of linguistic properties. Based on these properties and on their impact on inference, I furthermore proposed a classification of adjectives that establishes a set of 42 classes.

In this chapter, I make the inferential impact of this classification more precise by specifying for some of the defined classes, a first-order logic (FOL) semantic representation[34]. Combined with a small grammar fragment capturing the basic syntactic contexts in which adjectives can occur, this FOL semantics of adjectives not only gives a more precise model theoretic definition of the defined classes but also permits testing the predictions being made by the classification.

Section 8.1 begins by presenting the basic grammar fragments assumed for testing. Section 8.2 then specifies the semantics associated with the classes. Finally, section 8.3 briefly reports on the proof of concept implementation used to verify the predictions made by the grammar resulting from putting together the basic grammar fragment presented in Section 8.1 and the classes semantics proposed in Section 8.2.

8.1 The Representation Language

As has often been observed, not all of natural language meaning can be represented by first order logic. There are expressions such as, *most, I didn't, former, possible* whose meaning intuitively involve higher-order constructs.

[34]As we shall see, not all classes are assigned a specific semantic representation. Lexical semantics properties such as antonymy and stability/dynamicity are not captured directly in the formulae. This for two reasons, first differences due to antonymy are better accounted for by axioms as the distribution of antonyms in the different classes is not regular, i.e. it is not necessary the case that the set of antonyms of the members of an adjectival class are grouped in the same class. Second, capturing differences in stability/dynamicity would require a proper treatment of temporal semantics which lies outside the scope of this work.

Nevertheless, as Hobbs (1985) and others have argued, semantic representations for natural language need not be higher-order in that ontological promiscuity can help address the problem. That is, by reifying all objects that can be predicated of, it is possible to retain a semantic representation scheme for NL that is first-order. This observation is crucial for computational applications for two reasons. First, logics that go beyond first order are highly undecidable. Second and more importantly, there is no off-the-shelf higher order automated reasoner that could be used to reason about the meaning of higher-order formulae.

Following Hobbs (1985), I adopt a promiscuous ontology and assume that for every predication that can be made in natural language, there is a corresponding *eventuality*. As Hobbs has argued, this allows for higher order predications to remain first order in that they become predications over (first order) eventualities.
Thus, in the domain there are entities which are either *eventualities* or *individuals*. Moreover, like Hobbs, I assume a model to describe a platonic universe containing everything that can be spoken about whether or not these things exist in the real world. To express existence in the real world, a special predicate (*Exists*) is introduced. I use the following notation:

- e_i, for `eventuality` variables,
- x_i, for `individuals`,
- capital letters, e.g. P, Q, K, for `properties` of individuals,

Figure 8.1 summarises the semantic representations assigned to the language fragment relevant to the treatment of adjectives presented in this chapter.

In the implementation, I tested this grammar fragment on a list of sentences having the following syntactic structures: [NP Cop A N], [NP, Cop, A] and [NP, Cop, NP].
More specifically, I tested whether, given the semantic representation in Figure 8.1 and the semantic representation assigned to adjectives which will be discussed in the next sections, the following inferences hold for each class of adjectives:

$$AN \rightarrow A$$
$$AN \rightarrow N$$
$$AN \rightarrow \neg N$$

The sentences which were tested are listed for each class in appendix A and correspond to the patterns P31, P34 and P37.

8.1.1 The Semantics of Determiners

The semantic representation of determiners[35] is the usual one, except that it introduces an existential quantification over an eventuality variable e to which the noun property P is applied:

(185) **a/this:** $\lambda P \lambda Q \exists x \exists e.[P(e)(x) \wedge Q(x)]$

[35]For reason of simplicity and to reduce scope ambiguities, I will not account here for the determiner *every*.

Determiners:	
a/this	$\lambda P \lambda Q \exists x \exists e.[P(e)(x) \wedge Q(x)]$
Pronouns/NP:	
this/someone/something	$\lambda P \exists x.P(x)$
Nouns:	
cn	$\lambda Pol \lambda e \lambda x.[Pol(N'(e)) \wedge e = x]$
pn	$\lambda P.P(\text{pn})$
Verbs:	
iv	$\lambda x \exists e.[V(e) \wedge \text{arg}_{\mathit{1}}(e,x)]$
tv	$\lambda K \lambda x \exists e.K(\lambda y[V(e) \wedge \text{arg}_{\mathit{1}}(e,x) \wedge \text{arg}_{\mathit{2}}(e,y)])$
Copula:	
cop1	$\lambda K \lambda x.K(\lambda y.[x = y])$
cop1 neg	$\lambda K \lambda x.K(\lambda y.\neg[x = y])$
cop2	$\lambda K \lambda x \exists z.K(\lambda Pol \lambda e \lambda y.[x = z])(\lambda S.S)(\text{true})(z)$
cop2 neg	$\lambda K \lambda x \exists z.K(\lambda Pol \lambda e \lambda y.[x = z])(\lambda S.\neg S)(\text{true})(z)$
Auxiliaries:	
do/does	$\lambda V.V$
do/does not	$\lambda V \lambda z \exists y.[V(y) \wedge \neg[z = y]]$

Figure 8.1: Language Fragment

8.1.2 The Semantics of Nouns

The semantics of nouns reflect their possible interactions with the different types of adjectives (N stands for the property denoted by the noun).

(186) **noun:** $\lambda Pol \lambda e \lambda x.[Pol(N(e)) \wedge e = x]$

The additional lambda variable e is imposed by the treatment of adjective semantics I propose and more specifically, by the necessity to sometimes distinguish between the individual described by the noun N and the individual described by the adjective. The variable Pol accounts for the polarity of the noun, i.e. whether it occurs with a negation or not. The argument of the Pol variable is supplied by semantic construction rules[36]. The semantics of the pronouns *this/someone/something* which will be used in the derivations throughout this chapter is given in (188).

(188) **this/someone/something:** $\lambda P \exists x.P(x)$

[36]The following construction rules are used to control the scope of the negation.

(187) $Det + N \rightarrow Det(N(\lambda S.S))$
not $+ Det + N \rightarrow Det(N(\lambda S.\neg S))$

where N is noun or a noun modified by an adjective, i.e. $[A^*N]$. The example derivations in (194) shows what their effect is: it is the noun property which is negated and not the equivalence relation between the variable x and the noun concept.

8.1.3 The Semantics of Verbs

I assign to verbs a neodavidsonian (Parson, 1995) representation as in (189), where e is the eventuality introduced by the verb V and arg_i are binary relations which account for the semantic arguments of the eventuality e (V stands for the relation denoted by the verb).

(189) **iv:** $\lambda x \exists e.[V(e) \wedge \textit{arg}_1(e,x)]$
tv: $\lambda K \lambda x \exists e.K(\lambda y[V(e) \wedge \textit{arg}_1(e,x) \wedge \textit{arg}_2(e,y)])$

Further, the semantics of auxiliaries is represented as shown in (190).

(190) **do/does:** $\lambda V.V$
do/does not: $\lambda V \lambda z \exists y.[V(y) \wedge \neg[z=y]]$

An example of semantic constructions which use these representations is shown in (191).

(191) a. Mia smokes a cigarette
- *Mia (smokes (a cigarette))*
- *Mia (smokes (*$\lambda Q \exists x \exists e.[\textit{cigarette}(e) \wedge e = x \wedge Q(x)]$*))*
- *Mia (*$\lambda z \exists e_1 \exists x \exists e.[\textit{cigarette}(e) \wedge e = x \wedge \textit{smoke}(e_1) \wedge \textit{arg}_1(e_1,z) \wedge \textit{arg}_2(e_1,x)]$*)*
- $\exists e_1 \exists x \exists e.[\textit{cigarette}(e) \wedge e = x \wedge \textit{smoke}(e_1) \wedge \textit{arg}_1(e_1,mia) \wedge \textit{arg}_2(e_1,x)]$

b. Mia does not smoke a cigarette.
- *Mia (does not (smoke (a cigarette)))*
- *Mia (does not (*$\lambda z \exists e_1 \exists x \exists e.[\textit{cigarette}(e) \wedge e = x \wedge \textit{smoke}(e_1) \wedge \textit{arg}_1(e_1,z) \wedge \textit{arg}_2(e_1,x)]$*))*
- *Mia (*$\lambda z \exists y \exists e_1 \exists x \exists e.[\textit{cigarette}(e) \wedge e = x \wedge \textit{smoke}(e_1) \wedge \textit{arg}_1(e_1,y) \wedge \textit{arg}_2(e_1,x) \wedge \neg(z=y)]$*)*
- $\exists y \exists e_1 \exists x \exists e.[\textit{cigarette}(e) \wedge e = x \wedge \textit{smoke}(e_1) \wedge \textit{arg}_1(e_1,y) \wedge \textit{arg}_2(e_1,x) \wedge \neg(mia = y)]$

8.1.4 The Semantics of the Copula

Following Montague (1973), I assign a unique representation for both the uses of the copula in nominal identity statements (e.g. *John is Mary* → john=mary) and in nominal predicative assertions (e.g. *John is a man* → man(john)). The *cop2* representation, accounts for predicative constructions (e.g. *John is brave*) in which the predicate is an adjective. In this case, the type of the adjective needs to be adjusted to that of a noun phrase (i.e. $\lambda Pol \lambda e \lambda y.[x = z]$). Further the polarity variable which controls the scope of the negation in the decompositional representation of the semantics of the adjectives is instantiated with the values $\lambda S.\neg S$ or $\lambda S.S$, depending on whether the copula is negated or not.

(192) **cop1:** $\lambda K \lambda x.K(\lambda y.[x=y])$
cop2: $\lambda K \lambda x \exists z.K(\lambda Pol \lambda e \lambda y.[x=z])(\lambda S.S)(\textit{true})(z)$

(193) **cop1 neg:** $\lambda K \lambda x.K(\lambda y.\neg[x = y])$
cop2 neg: $\lambda K \lambda x \exists z.K(\lambda Pol \lambda e \lambda y.[x = z])(\lambda S.\neg S)(\textit{true})(z)$

An example of semantic construction based on this semantic representation is given in (194).

(194) a. Mia is a woman
- *Mia (is (a woman))*
- *Mia is* $(\lambda P \lambda Q \exists x \exists e.[P(e)(x) \wedge Q(x)]((\textit{woman})(\lambda S.S)))$
- *Mia (is* $(\lambda Q \exists x \exists e.[\textit{woman}(e) \wedge e = x \wedge Q(x)]))$
- *Mia* $(\lambda z \exists x \exists e.[\textit{woman}(e) \wedge e = x \wedge x = z])$
- $\exists x \exists e.[\textit{woman}(e) \wedge e = x \wedge x = mia]$

b. Mia is not a man.
- *Mia (is not (a man))*
- *Mia is* $(\lambda P \lambda Q \exists x \exists e.[P(e)(x) \wedge Q(x)]((\textit{man})(\lambda S.\neg S)))$
- *Mia (is not* $(\lambda Q \exists x \exists e.[\neg \textit{man}(e) \wedge e = x \wedge Q(x)]))$
- *Mia* $(\lambda z \exists x \exists e.[\neg \textit{man}(e) \wedge e = x \wedge x = z])$
- $\exists x \exists e.[\neg \textit{man}(e) \wedge e = x \wedge x = mia]$

In the following section, I show how this fragment can be extended with a semantics for adjectives which models some of the distinctions captured in the classification I proposed in the previous chapter.

8.2 Semantics of Adjectives

As I have shown in the previous chapter, the semantics of [A N]$_{AP}$ phrases has very different inferential properties depending on the type of the adjective A. These differences reflect the interactions between model theoretic, lexical semantic, syntactic and morphoderivational properties of adjectives and can be summarised in the following three main points:

The properties licensed by the adjective and the noun to contribute to the meaning of the [A N]$_{AP}$ phrase.
Depending on the adjective type, the properties denoted by A and N will contribute either directly or indirectly to the meaning of the [A N]$_{AP}$ phrase. Thus in an intersective [A N]$_{AP}$ phrase, the meaning contributed by A and N are simply the properties they denote. By contrast, the privative *fake* forces the negation of the N property to be part of the [A N]$_{AP}$ meaning whereas the subsective *gastronomical* induces a relation to the morphoderivationally related noun concept (*about gastronomy*) to be included in the [A N]$_{AP}$ meaning. More generally, the properties that compose the meaning of the [A N]$_{AP}$ phrase can be the denotation of A and/or N, the negation of N, its denotation in the past or in the future or some other property derived from them.

The existence in the real world of the entity denoted by the NP.
In all cases the $[A\ N]_{AP}$ phrase denotes a set of individuals but whilst in most cases the $[A\ N]_{AP}$ phrase is neutral with respect to the real world existence of these individuals, some privatives (e.g. *fictitious*) induce $[A\ N]_{AP}$ phrases to explicitly negate it (a *fictitious friend* does not exist in the real world).

The number of individuals introduced by the [A N]$_{AP}$ phrase.
Thus, *the red table* evokes a single individual x which is both *red* and a *table* whereas *the gastronomical book* refers to a book x which is about the gastronomy concept y. More generally, the variables predicated of by the noun and by the adjective can refer either to the same or to two distinct individual(s).

I assign to all adjectives the general semantic representation scheme in (196), where A' represents the property licensed by the adjective. R_1 and R_2 accounts for the relations between the individual/eventuality introduced by the adjective and that introduced by the noun. The variable n ranges over noun constants and is used to block illicit inference about the adjectival properties and to access lexical semantic axioms describing modifying relations between adjective and noun, (e.g. the interactions between adjectives of different classes and nouns with regards to transparency[37]). N' represents the property denoted by the noun. Pol and Pol_n are polarity arguments of value either $\lambda S.S$ or $\lambda S.\neg S$. The value of the variable Pol_n depends on the semantic class of the adjective (e.g. privatives such as *fake* instantiate the value of Pol_n to $\lambda S.\neg S$). The variable Pol will be applied differently, depending on the decompositional semantics of the adjective. Finally, the lambda variable e is introduced to adjust the result of the adjective noun application to that of a noun.

(196) $\lambda N' \lambda Pol \lambda e \lambda x \exists e_a \exists e_n.[Pol(A'(e_a)) \wedge R_1(e_a, x) \wedge R_2(e_a, n) \wedge N'(Pol_n)(e_n)(x)]$

This representation captures the previous observations as follows. First, the meaning of the $[A\ N]_{AP}$ phrase is a function not of the A and N meaning but rather of properties derived from these meanings (A' for A and N' for N). For instance, the variable Pol_n permits privative adjectives to introduce the negation of N. This accounts for the first observation.
Second, A' the particular property licensed by the adjective will sometimes license the use of the *Exists* predicate, thereby permitting to distinguish between existence in the universe of discourse and existence in the real world of the individuals denoted by the

[37] As an example, consider the representations assigned to the sentences (195) which will be discussed in section 8.2.2, where $R2$ is instantiated to *role_as*(e_a,n). In this case, the different instantiations of the variable n (e.g. *king* and *john*) disallow to infer (195b) from (195a).

(195) a. *John is a cruel king*
$cruelty(e_a) \wedge \mathit{theme}(e_a, x) \wedge role\_as(e_a, king) \wedge \mathit{king}(e_n) \wedge e_n = john$

b. *John is cruel*
$cruelty(e_a) \wedge \mathit{theme}(e_a, x) \wedge role\_as(e_a, john) \wedge e_n = john$

[A N]$_{AP}$ phrase. For instance, in the case of privatives such as *fictitious*, the decompositional meaning of A' will include the negation of this predicate, $\neg \textsf{\textit{Exist}}(x)$.
Third, it introduces an existential quantification (in the platonic universe) over not one but two variables (e_a and e_n). Depending on how the formula is instantiated (and in particular on the value of R_1 and R_2) these two variables may or may not denote the same object. This accounts for the third observation according to which an [A N]$_{AP}$ phrase may refer to either one or two individuals.

I now show how this general scheme receives different instantiations depending on the adjectival class being considered and how each instantiation, together with the set of relevant/appropriate lexical axioms, predicts the inferential patterns expected for each class. The idea is to decompose the meaning of adjectives into finer grained lexical meanings so as to account for the interactions between the different levels of linguistic description. Depending on the lexical meaning involved, this decomposition induces different instantiation patterns for the A', N' and for the relations R_1 and R_2 mentioned in the general scheme for adjective semantic representation.
In the following, for ease of description I will sometimes give an abbreviate representation of the semantics of adjectives without lambda variables.

8.2.1 Intersective Adjectives

For intersective adjectives, the nouns and adjective predications hold of the same individual. Moreover, the nominal predication is asserted rather than negated. Accordingly, the general schema assigned to intersective specialises the schema for adjectives given in (196) as follows :

(197) $\lambda N' \lambda Pol \lambda e \lambda x \exists e_a \exists e_n.[Pol(\mathsf{A}'(e_a)) \wedge R_1(e_a, x) \wedge N'(\lambda S.S)(e_n)(x)]$

That is, the noun polarity Pol_n is instantiated to $\lambda S.S$ and R_2 to the empty relation $\lambda x \lambda y.true$.

Different instantiations of the R_1 relation then make it possible to distinguish between the various subclasses identified in the previous chapter as shown in Figure 8.2. We now discuss in more detail the various instantiations proposed for the 4 main classes of intersectives namely, (i) common intersectives (e.g. Ipa7, Ipa9, Ipa10, Ipa12), (ii) the gradable intersectives (e.g. Ipa1, Ipa2, Ipa3, Ipa4, Ipa5, Ip1), (iii) the noun-related intersectives (e.g. Ipa6, Ipa8, Ia1) and (iv) the verb-related intersectives (e.g. Ipa11, Ip2).

Common Intersectives

For common intersectives, the decompositional meaning of A' corresponds to that of the adjective A and the relation R_1 to the identity relation (e.g. $e_a = x$). For instance, the semantics of *oval* specialises the representation schema given in (197) as follows:

[38]In the Figures of this chapter, all free variables which appear in the formulae are considered bound by $\lambda N' \lambda Pol \lambda e \lambda x$. Further, n_a, v_a and adj are variable which, in the same way as n, range respectively

Intersective Adjectives[38]	
Ipa7	$\exists e_a \exists e_n [Pol(\mathsf{A}(e_a)) \wedge e_a = x \wedge N'(\lambda S.S)(e_n)(x)]$ e.g. *oval*
Ipa9	$\exists e_a \exists e_n [Pol(\mathsf{A}(e_a)) \wedge e_a = x \wedge N'(\lambda S.S)(e_n)(x)]$ e.g. *finite*
Ipa10	$\exists e_a \exists e_n [Pol(\mathsf{A}(e_a)) \wedge e_a = x \wedge N'(\lambda S.S)(e_n)(x)]$ e.g. *genuine*
Ipa12	$\exists e_a \exists e_n [Pol(\mathsf{A}(e_a)) \wedge e_a = x \wedge N'(\lambda S.S)(e_n)(x)]$ e.g. *absent*
Ipa1	$\exists e_a \exists e_n \exists d [Pol(\mathsf{N}_a(e_a)) \wedge \textsf{extent}(e_a, d) \wedge \textsf{has-p}(e_a, x) \wedge N'(\lambda S.S)(e_n)(x)]$ e.g. *red*
Ipa2	$\exists e_a \exists e_n \exists a \exists d [\mathsf{N}_a(e_a) \wedge \textsf{extent}(e_a, d) \wedge \textsf{absStandard}(a, n_a) \wedge Pol(\textsf{low/high}(d, a)) \wedge \textsf{has-p}(e_a, x) \wedge N'(\lambda S.S)(e_n)(x)]$ e.g. *wet*
Ipa3	$\exists e_a \exists e_n \exists a \exists d [\mathsf{N}_a(e_a) \wedge \textsf{extent}(e_a, d) \wedge \textsf{absStandard}(a, n_a) \wedge Pol(\textsf{low/high}(d, a)) \wedge \textsf{has-p}(e_a, x) \wedge N'(\lambda S.S)(e_n)(x)]$ e.g. *straight*
Ipa4	$\exists e_a \exists e_n \exists a \exists d [\mathsf{N}_a(e_a) \wedge \textsf{extent}(e_a, d) \wedge \textsf{absStandard}(a, n_a) \wedge Pol(\textsf{low/high}(d, a)) \wedge \textsf{has-p}(e_a, x) \wedge N'(\lambda S.S)(e_n)(x)]$ e.g. *spicy*
Ipa5	$\exists e_a \exists e_n \exists a \exists d [\mathsf{N}_a(e_a) \wedge \textsf{extent}(e_a, d) \wedge \textsf{absStandard}(a, n_a) \wedge Pol(\textsf{low/high}(d, a)) \wedge \textsf{has-p}(e_a, x) \wedge N'(\lambda S.S)(e_n)(x)]$ e.g. *clean*
Ip1	$\exists e_a \exists e_n \exists a \exists d [\mathsf{V}_a(e_a) \wedge \textsf{extent}(e_a, d) \wedge \textsf{absStandard}(a, v_a) \wedge Pol(\textsf{theme}(e_a, x)) \wedge N'(\lambda S.S)(e_n)(x)]$ e.g. *asleep*
Ipa6	$\exists e_a \exists e_n [Pol(\mathsf{N}_a(e_a)) \wedge e_a = x \wedge N'(\lambda S.S)(e_n)(x)]$ e.g. *carnivorous*
Ipa8	$\exists e_a \exists e_n [Pol(\mathsf{N}_a(e_a)) \wedge \textsf{Rel}(e_a, x) \wedge N'(\lambda S.S)(e_n)(x)]$ e.g. *artisanal*
Ia1	$\exists e_a \exists e_n [Pol(\mathsf{A}(e_a)) \wedge \textsf{time}(e_a, x) \wedge N'(\lambda S.S)(e_n)(x)]$ e.g. *present*
Ipa11	$\exists e_a \exists e_n [\mathsf{V}_a(e_a) \wedge Pol(\textsf{patient}(e_a, x)) \wedge N'(\lambda S.S)(e_n)(x)]$ e.g. *closed*
Ip2	$\exists e_a \exists e_n [\mathsf{V}_a(e_a) \wedge Pol(\textsf{theme}(e_a, x)) \wedge N'(\lambda S.S)(e_n)(x)]$ e.g. *afloat*

Figure 8.2: Semantics of Intersective Adjectives

(198) **oval:** $\lambda N'\lambda Pol\lambda e\lambda x\exists e_a\exists e_n.[Pol(\textsf{\textit{oval}}(e_a)) \wedge e_a = x \wedge N'(\lambda S.S)(e_n)(x)]$

Based on this representation, (199) shows the semantic construction for the sentence *Something is an oval table*.

(199)

- *Something (is (an (oval table)))*
- *Something is (an (oval* $(\lambda Pol\lambda e\lambda x[Pol(\textsf{\textit{table}}(e)) \wedge x = e]$*)))*
- *Something is (an* $(\lambda N'\lambda Pol\lambda e\lambda x\exists e_a\exists e_n.[Pol(\textsf{\textit{oval}}(e_a)) \wedge e_a = x \wedge N'(\lambda S.S)(e_n)(x)]$ $(\lambda Pol\lambda e\lambda x[Pol(\textsf{\textit{table}}(e)) \wedge x = e]$*)))*
- *Something is (an ((*$\lambda Pol\lambda e\lambda x[\textsf{\textit{oval}}(e_a) \wedge e_a = x \wedge \textsf{\textit{table}}(e_n) \wedge e_n = x]$*)(*$\lambda S.S$*)))*
- *Something is (*$\lambda Q\exists x\exists e\exists e_a\exists e_n[\textsf{\textit{oval}}(e_a) \wedge e_a = x \wedge \textsf{\textit{table}}(e_n) \wedge e_n = x \wedge Q(x)]$*)*
- *Something (*$\lambda z\exists x\exists e\exists e_a\exists e_n[\textsf{\textit{oval}}(e_a) \wedge e_a = x \wedge \textsf{\textit{table}}(e_n) \wedge e_n = x \wedge x = z]$*)*
- $\exists y\exists x\exists e\exists e_a\exists e_n[\textsf{\textit{oval}}(e_a) \wedge e_a = x \wedge \textsf{\textit{table}}(e_n) \wedge e_n = x \wedge x = y]$*)*

This representation can be simplified to the formula (200):

(200) $\exists x\exists e_a\exists e_n.[\textsf{\textit{oval}}(e_a) \wedge e_a = x \wedge \textsf{\textit{table}}(e_n) \wedge e_n = x]$

This correctly entails that there is an entity x which is both oval and a table, i.e.

(200)	$\models$	$\exists x\exists e_a.[\textsf{\textit{oval}}(e_a) \wedge e_a = x]$	*x is oval*
(200)	$\models$	$\exists x\exists e_n.[\textsf{\textit{table}}(e_n) \wedge e_n = x]$	*x is a table*

Although I shall not do so here, the four subclasses of common intersectives can be further distinguished by means of lexical axioms accounting for the different antonymic relations each subclass involves (e.g. multi-opposition for Ipa7 *vs.* binary opposition for Ipa9, Ipa10 and Ipa12). Syntactic differences can be accounted for in the grammar in order to distinguish between the different complements the classes allow (e.g. Ipa10 allows subject embedding whereas Ipa9 and Ipa12 do not). And differences in dynamicity/stability can be accounted for by extending the approach with a adequate treatment of temporal semantics.

Gradable Intersectives

Following a long tradition of work on this topic (Kennedy, 2004; Kennedy and McNally, 2005; Kennedy, 2005; Schwarzschild and Wilkinson, 2002; Klein, 1991; Bierwisch, 1989; Heim, 1985; von Stechow, 1984), I assume that a gradable adjective (e.g. *tall* and *wet*) expresses a relation between a degree d and an individual *x*. This presumes that every gradable adjective includes as part of its meaning a measure function: a function from individuals to degrees on a scale.

over noun, verb or adjective constants.

By observing that in WordNet these adjectives are semantically related to nouns (N_a) denoting properties of which they are attributes of (e.g. *height/wetness*), it is possible to specify an appropriate decompositional semantics for A'.
The subset of gradable intersective adjectives denote gradable properties with absolute standard (e.g. *dry*, *ill*, *wet*). Such adjectives can be assigned a decompositional semantics A' as in (201), where $\textsf{extent}$ accounts for the value (d) of the scalar property N_a described by the adjective, which is compared with the value of the context independent absolute standard a of the property N_a individuated by the adjective.

(201) $A' : \exists a \exists d[N_a(e_a) \wedge \textsf{extent}(e_a, d) \wedge (\neg)\textsf{low/high}(d, a) \wedge \textsf{absStandard}(a, N_a)]$

Further we say that R_1 denotes a *has-p* (for "has property") relation between the eventuality e_a predicated of by the adjective and the individual x.

For instance, the semantic representation assigned to the adjectives *dry* and *wet* is given in (202).

(202)

dry: $\lambda N' \lambda Pol \lambda e \lambda x \exists e_a \exists e_n \exists a \exists d[\textsf{wetness}(e_a) \wedge \textsf{extent}(e_a, d) \wedge \textsf{absStandard}(a, \textsf{wetness}) \wedge Pol(\textsf{low}(d, a)) \wedge \textsf{has-p}(e_a, x) \wedge N'(\lambda S.S)(e_n)(x)]$

wet: $\lambda N' \lambda Pol \lambda e \lambda x \exists e_a \exists e_n \exists a \exists d.[\textsf{wetness}(e_a) \wedge \textsf{extent}(e_a, d) \wedge \textsf{absStandard}(a, \textsf{wetness}) \wedge Pol(\neg\textsf{low}(d, a)) \wedge \textsf{has-p}(e_a, x) \wedge N'(\lambda S.S)(e_n)(x)]$

Thus, the semantic representation of *Something is a dry dishcloth* given in (203)

(203) $\exists x \exists e_a \exists e_n \exists a \exists d.[\textsf{wetness}(e_a) \wedge \textsf{extent}(e_a, d) \wedge \textsf{absStandard}(a, wetness) \wedge \textsf{low}(d, a) \wedge \textsf{has-p}(e_a, x) \wedge \textsf{dishcloth}(e_n) \wedge e_n = x]$

correctly entails that there is an entity x which is both dry and a dishcloth and that has a low degree of wetness, i.e.

(203) $\models \exists x \exists e_a \exists a \exists d.[\textsf{wetness}(e_a) \wedge \textsf{extent}(e_a, d) \wedge \textsf{absStandard}(a, \textsf{wetness}) \wedge \textsf{low}(d, a) \wedge \textsf{has-p}(e_a, x)]$ *x is dry*

(203) $\models \exists x \exists e_n.[\textsf{dishcloth}(e_n) \wedge e_n = x]$ *x is a dishcloth*

(203) $\models \exists x \exists e_a \exists a \exists d.[\textsf{wetness}(e_a) \wedge \textsf{extent}(e_a, d) \wedge \textsf{absStandard}(a, \textsf{wetness}) \wedge \neg\neg\textsf{low}(d, a) \wedge \textsf{has-p}(e_a, x)]$ *x is not wet*

Gradable intersective adjectives further differ in their behaviour with respect to antonymy. The adjectives in the Ipa1 class enter in multi-opposition relation with their antonyms, whereas the adjectives in the other subclasses of gradable intersectives enter in binary opposition relations. This difference can be accounted for by means of lexical axioms.
A further distinction can be made between the class Ip1 which groups adjectives semantically related to verbs and the classes Ipa2, Ipa3, Ipa4 and Ipa5 which do not.
I do not account for differences in stability/dynamicity so, that the classes Ipa2, Ipa3 Ipa4 and Ipa5 have the same semantic representation.

Noun-related Intersectives

The lexical semantics of noun-related intersective adjectives can be captured by instantiating the meaning of the adjective A' directly with the meaning of the related noun N_a, as in the case of the class Ipa6, or with a more complex relation based on the meaning of N_a as in the case of the classes Ipa8 and Ipa1.
For instance, the semantics of adjectives such as *carnivorous, liquid* (Ipa6) can be captured by instantiating A' with the corresponding noun property N_a and R_1 with the identity relation, i.e. $e_a = x$.
As an example, the semantic representation assigned to the adjective *liquid* is a s given in (204)

(204) **liquid:** $\lambda N' \lambda Pol \lambda e \lambda x \exists e_a \exists e_n .[Pol(\textsf{\textit{liquid}}(e_a)) \wedge e_a = x \wedge N'(\lambda S.S)(e_n)(x)]$

Given this representation, the semantics of *Something is a liquid substance* is as in (205)

(205) $\exists x \exists e_a \exists e_n.[\textsf{\textit{liquid}}(e_a) \wedge e_a = x \wedge \textsf{\textit{substance}}(e_n) \wedge e_n = x]$

and correctly entails that there is an entity x which is both a liquid and a substance, i.e.

(205)	$\models$	$\exists x \exists e_a.[\textsf{\textit{liquid}}(e_a) \wedge e_a = x]$	*x is liquid*
(205)	$\models$	$\exists x \exists e_a.[\textsf{\textit{liquid}}(e_a) \wedge e_a = x]$	*x is a liquid*
(205)	$\models$	$\exists x \exists e_n.[\textsf{\textit{substance}}(e_n) \wedge e_n = x]$	*x is a substance*

For adjectives belonging to the Ipa8 class (e.g. *weekly, artisanal*) the relation between the individuals and the eventuality introduced by the related noun is more complex and has a different value depending on the taxonomical category of the adjective. As an example, the representation for the adjective *weekly* is given in (206).

(206) **weekly:** $\lambda N' \lambda Pol \lambda e \lambda x \exists e_a \exists e_n \exists d.[\textsf{\textit{frequency}}(e_a) \wedge \textsf{\textit{week}}(d) \wedge \textsf{\textit{value}}(e_a, d) \wedge Pol(\textsf{\textit{has-p}}(e_a, x)) \wedge N'(\lambda S.S)(e_n)(x)]$

Deverbal Intersectives

Finally, the lexical semantics of intersective adjectives which are related to intransitive verbs can be captured by instantiating A' with the semantically related verb V_a and R_1 with the theta role relation imposed by the particular verb class the adjective is related to.
As an example (207) shows the semantics assigned to the adjectives *afloat* and *aground* belonging to the Ip2 class and to the adjective *closed* member of the Ipa11 class. The difference between this classes is at ontological level. The verbs in the Ipa11 class are transitive and affect their object in some way (thus to the modified individual x is assigned the patient role), whereas the verbs in Ip2 are intransitive (thus to the modified individual x is assigned a theme role).

(207)

afloat: $\lambda N'\lambda Pol\lambda e\lambda x\exists e_a\exists e_n\,.[\textit{float}(e_a)\wedge Pol(\textit{theme}(e_a,x))\wedge N'(\lambda S.S)(e_n)(x)]$
aground: $\lambda N'\lambda Pol\lambda e\lambda x\exists e_a\exists e_n.[\textit{float}(e_a)\wedge Pol(\neg\textit{theme}(e_a,x))\wedge N'(\lambda S.S)(e_n)(x)]$
closed: $\lambda N'\lambda Pol\lambda e\lambda x\exists e_a\exists e_n.[\textit{close}(e_a)\wedge Pol(\textit{patient}(e_a,x))\wedge N'(\lambda S.S)(e_n)(x)]$

Thus, the semantic representation of *A boat is afloat* in (208)

(208) $\exists x\exists e_a\exists e_n.[\textit{float}(e_a)\wedge\textit{theme}(e_a,x)\wedge\textit{boat}(e_n)\wedge e_n=x]$

allows the same entailment patterns as the intransitive verb *float*, i.e.

(208) $\models$ $\exists x\exists e_n.[\textit{boat}(e_n)\wedge e_n=x]$ *x is a boat*
(208) $\models$ $\exists x\exists e_a.[\textit{float}(e_a)\wedge\textit{theme}(e_a,x)]$ *x is afloat*
(208) $\models$ $\exists x\exists e_a\exists e_n.[\textit{float}(e_a)\wedge\textit{theme}(e_a,x)\wedge\textit{boat}(e_n)\wedge e_n=x]$ *A boat floats*
(208) $\models$ $\exists x\exists e_a\exists e_n.[\textit{float}(e_a)\wedge\textit{theme}(e_a,x)\wedge\textit{boat}(e_n)\wedge e_n=x]$ *A boat is the floater*
(208) $\models$ $\exists x\exists e_a\exists e_n.[\textit{float}(e_a)\wedge\neg\neg\textit{theme}(e_a,x)\wedge\textit{boat}(e_n)\wedge e_n=x]$ *A boat is not aground*

8.2.2 Subsective Adjectives

Subsective adjectives (e.g. *big*, *small*) are characterised by the fact that the [A N]$_{AP}$ phrase entails N but not A.
Thus in such cases, both R_1 and R_2, the modifying relation between adjective and noun, are not empty. Depending on the adjectival class considered, R1 is instantiated to the has-p relation or to some relation sem-$role$ denoting a thematic or a semantic role (e.g. *theme*, *activity*). The relation R2 accounts for relations with the context dependent standard of the property denoted by the adjective (e.g. $standard$) in the case of dimensional subsectives or is instantiated with $role$-as relation which is used to disallow illicit inference.

The instantiations of the general scheme for adjectival semantics for the 19 subclasses of subsectives is given in Figure 8.3.
As an example, I describe the representations assigned to the following subclasses, namely (i) gradable dimensional subsective adjectives (e.g. Spa2, Spa3, Spa4, Spa5), (i) nondimensional subsectives (e.g. Spa6-Spa14, Sp1) and (iii). relational subsectives (e.g. Sa1).

Dimensional Subsectives

Dimensional subsectives include gradable adjectives with context sensitive standard. The semantics assigned to such adjectives follows the general scheme given in (209), in which the *extent* accounts for the value d of the scalar property N_a described by the adjective, which is compared with the context dependent average value a of the standard individuated by the noun property N' (i.e. n) the adjective modifies.

Subsective Adjectives		
Spa1	$\exists e_a \exists e_n \exists a \exists d[\mathsf{N}_a(e_a) \wedge Pol(\textsf{has-p}(e_a, x)) \wedge \textsf{extent}(e_a, d) \wedge \textsf{standard}(a, n) \wedge Pol(\textsf{max/min/average}(d, a)) \wedge N'(\lambda S.S)(e_n)(x))]$	e.g. *enormous*
Spa2	$\exists e_a \exists e_n \exists a \exists d[\mathsf{N}_a(e_a) \wedge Pol(\textsf{has-p}(e_a, x)) \wedge \textsf{extent}(e_a, d) \wedge Pol(\textsf{highFor}(d, a)) \wedge \textsf{standard}(a, n) \wedge \wedge N'(\lambda S.S)(e_n)(x))]$	e.g. *big*
Spa3	$\exists e_a \exists e_n \exists a \exists d[\mathsf{N}_a(e_a) \wedge Pol(\textsf{has-p}(e_a, x)) \wedge \textsf{extent}(e_a, d) \wedge Pol(\textsf{lowFor}(d, a)) \wedge \textsf{standard}(a, n) \wedge N'(\lambda S.S)(e_n)(x))]$	e.g. *small*
Spa4	$\exists e_a \exists e_n \exists a \exists d[\mathsf{N}_a(e_a) \wedge Pol(\textsf{has-p}(e_a, x)) \wedge \textsf{extent}(e_a, d) \wedge Pol(\textsf{highFor}(d, a)) \wedge \textsf{standard}(a, n) \wedge \wedge N'(\lambda S.S)(e_n)(x))]$	e.g. *fast*
Spa5	$\exists e_a \exists e_n \exists a \exists d[\mathsf{N}_a(e_a) \wedge Pol(\textsf{has-p}(e_a, x)) \wedge \textsf{extent}(e_a, d) \wedge Pol(\textsf{lowFor}(d, a)) \wedge \textsf{standard}(a, n) \wedge N'(\lambda S.S)(e_n)(x))]$	e.g. *slow*
Spa6	$e_n \exists e_a \exists e_n[\mathsf{N}_a(e_a) \wedge Pol(\textsf{sem-role}(e_a, x)) \wedge \textsf{role\_as}(e_a, n) \wedge \textsf{N'}(\lambda S.S)(e_n)(x)]$	e.g. *dangerous*
Spa7	$\exists e_a \exists e_n[\mathsf{N}_a(e_a) \wedge Pol(\textsf{sem-role}(e_a, x)) \wedge \textsf{role\_as}(e_a, n) \wedge \textsf{N'}(\lambda S.S)(e_n)(x)]$	e.g. *certain*
Spa8	$\exists e_a \exists e_n[\mathsf{N}_a(e_a) \wedge Pol(\textsf{sem-role}(e_a, x)) \wedge \textsf{role\_as}(e_a, n) \wedge \textsf{N'}(\lambda S.S)(e_n)(x)$	e.g. *necessary*
Spa9	$\exists e_a \exists e_n[\mathsf{N}_a(e_a) \wedge Pol(\textsf{theme}(e_a, x)) \wedge \textsf{role\_as}(e_a, n) \wedge \textsf{N'}(\lambda S.S)(e_n)(x)]$	e.g. *clear*
Spa10	$\exists e_a \exists e_n[\mathsf{N}_a(e_a) \wedge Pol(\textsf{experiencer}(e_a, x)) \wedge \textsf{role\_as}(e_a, n) \wedge \textsf{N'}(\lambda S.S)(e_n)(x)]$	e.g. *eager*
Spa11	$\exists e_a \exists e_n[\mathsf{N}_a(e_a) \wedge Pol(\textsf{theme}(e_a, x)) \wedge \textsf{role\_as}(e_a, n) \wedge \textsf{N'}(\lambda S.S)(e_n)(x)]$	e.g. *intelligent*
Spa12	$\exists e_a \exists e_n[\mathsf{N}_a(e_a) \wedge Pol(\textsf{activity}(e_a, e_n)) \wedge \textsf{role\_as}(e_a, n) \wedge \textsf{N'}(\lambda S.S)(e_n)(x)]$	e.g. *easy*
Spa13	$\exists e_a \exists e_n[\mathsf{N}_a(e_a) \wedge Pol(\textsf{theme}(e_a, x)) \wedge \textsf{role\_as}(e_a, n) \wedge \textsf{N'}(\lambda S.S)(e_n)(x)]$	e.g. *cruel*
Spa14	$\exists e_a \exists e_n[\mathsf{N}_a(e_a) \wedge Pol(\textsf{experiencer}(e_a, x)) \wedge \textsf{role\_as}(e_a, n) \wedge \textsf{N'}(\lambda S.S)(e_n)(x)]$	e.g. *sad*
Sp1	$\exists e_a \exists e_n[\mathsf{N}_a(e_a) \wedge Pol(\textsf{experiencer}(e_a, x)) \wedge \textsf{role\_as}(e_a, n) \wedge \textsf{N'}(\lambda S.S)(e_n)(x)]$	e.g. *averse*
Sa1	$\exists e_a \exists e_n[\mathsf{N}_a(e_a) \wedge Pol(\textsf{Rel}(e_a, e_n)) \wedge N'(\lambda S.S)(e_n)(x)]$	e.g. *gastronomical*
Sa2	$\exists e_a \exists e_n[\mathsf{N}_a(e_a) \wedge \textsf{ranking}(e_a, d) \wedge \textsf{has-p}(e_a, x) \wedge \textsf{role\_as}(d, n) \wedge Pol(\textsf{value}(d, adj)) \wedge \textsf{N'}(\lambda S.S)(e_n)(x)]$	e.g. *main*
Sa3	$\exists e_a \exists e_n[\mathsf{N}_a(e_a) \wedge \textsf{ranking}(e_a, d) \wedge \textsf{has-p}(e_a, x) \wedge \textsf{role\_as}(d, n) \wedge Pol(\textsf{value}(d, adj)) \wedge \textsf{N'}(\lambda S.S)(e_n)(x)]$	e.g. *utter*
Sa4	$\exists e_a \exists e_n[\mathsf{N}_a(e_a) \wedge \textsf{ranking}(e_a, d) \wedge \textsf{has-p}(e_a, x) \wedge \textsf{role\_as}(d, n) \wedge Pol(\textsf{value}(d, adj)) \wedge \textsf{N'}(\lambda S.S)(e_n)(x)]$	e.g. *second*

Figure 8.3: Semantics of Subsective Adjectives

(209) $\lambda N'\lambda Pol\lambda e\lambda x.\exists e_a\exists e_n\exists a\exists d[\textsf{N}_a(e_a)\wedge Pol(\textsf{has-p}(e_a,x))\wedge\textsf{extent}(e_a,d)\wedge\textsf{standard}(a,n)\wedge Pol(\textsf{low/high}(d,a))\wedge\wedge N'(\lambda S.S)(e_n)(x))]$

Given that the semantic representations assigned to the adjectives *tall* and *small* is as in (210),

(210)

tall: $\lambda N'\lambda Pol\lambda e\lambda x.\exists e_a\exists e_n\exists d\exists a[\textsf{size}(e_a)\wedge Pol(\textsf{has-p}(e_a,x))\wedge\textsf{extent}(e_a,d)\wedge Pol(\textsf{highFor}(d,a))\wedge\textsf{standard}(a,n)\wedge N'(\lambda S.S)(e_n)(x))]$

small: $\lambda N'\lambda Pol\lambda e\lambda x.\exists e_a\exists e_n\exists a\exists d[\textsf{size}(e_a)\wedge Pol(\textsf{has-p}(e_a,x))\wedge\textsf{extent}(e_a,d)\wedge Pol(\textsf{lowFor}(d,a))\wedge\textsf{standard}(a,n)\wedge N'(\lambda S.S)(e_n)(x))]$

the meaning of the adjectival phrase *Bixi is a big mouse* is

(211) $\exists x\exists e_a\exists e_n\exists a\exists d.[\textsf{size}(e_a)\wedge\textsf{has-p}(e_a,x)\wedge\textsf{extent}(e_a,d)\wedge\textsf{highFor}(d,a)\wedge\textsf{standard}(a,mouse)\wedge\textsf{mouse}(e_n)\wedge e_n=x\wedge x=\textsf{bixy}]$

which correctly entails that there is an entity x which is a mouse and has a big size relative to the average size of a mouse but not that x is big (in general), i.e.

(211) $\models\exists x\exists e_n.[\textsf{mouse}(e_n)\wedge e_n=x]$ *x is a mouse*

(211) $\models\exists x\exists e_a\exists a\exists d.[\textsf{size}(e_a)\wedge\textsf{extent}(e_a,d)\wedge\textsf{highFor}(d,a)\wedge\textsf{standard}(a,mouse)\wedge\textsf{has-p}(e_a,x)]$ *x has a big size relative to the average size of a mouse*

(211) $\not\models\exists x\exists e_a\exists a\exists d.[\textsf{size}(e_a)\wedge\textsf{extent}(e_a,d)\wedge\textsf{highFor}(d,a)\wedge\textsf{standard}(a,bixi)\wedge\textsf{has-p}(e_a,x)]$ *bixi is big*

Nondimensional Subsectives

The adjectives in this group (e.g. *cruel, eager, easy*) denote gradable nonscalar properties. They are related to nouns representing eventualities in which the individual denoted by the modified noun is involved as a role. Their semantics follows the general scheme given in (212).

(212) $\lambda N'\lambda Pol\lambda e\lambda x.\exists e_a\exists e_n[\textsf{N}_a(e_a)\wedge Pol(\textsf{sem-role}(e_a,x))\wedge\textsf{role\_as}(e_a,n)\wedge\textsf{N'}(\lambda S.S)(e_n)(x)]$

As an example, the semantics of the adjectives *eager, cruel* and *necessary* is given in (213).

(213)

eager: $\lambda N'\lambda Pol\lambda e\lambda x\exists e_a\exists e_n.[N_a(e_a)\wedge Pol(\textsf{experiencer}(e_a,x))\wedge\textsf{role\_as}(e_a,n)\wedge N'(\lambda S.S)(e_n)(x)]$

cruel: $\lambda N'\lambda Pol\lambda e\lambda x\exists e_a\exists e_n.[N_a(e_a)\wedge Pol(\textsf{theme}(e_a,x))\wedge\textsf{role\_as}(e_a,n)\wedge N'(\lambda S.S)(e_n)(x)]$

necessary: $\lambda N'\lambda Pol\lambda e\lambda x\exists e_a\exists e_n.[N_a(e_a)\wedge Pol(\textsf{topic}(e_a,x))\wedge\textsf{role\_as}(e_a,n)\wedge N'(\lambda S.S)(e_n)(x)]$

Thus the semantic representation assigned to *Vincent is a cruel husband*, i.e.

(214) $\exists x \exists y \exists e_a \exists e_n.[\textit{cruelty}(e_a) \wedge \textit{theme}(e_a, x) \wedge \textit{role\_as}(e_a, husband) \wedge \textit{husband}(e_n) \wedge e_n = x]$

correctly entails that Vincent is cruel in his role as a husband but not necessarily in other roles (e.g. as a friend) or in general (e.g. as a man), i.e.

(214) $\not\models \exists x \exists y \wedge \exists e_a \exists e_n.[\textit{cruelty}(e_a) \wedge \textit{theme}(e_a, x) \wedge \textit{role\_as}(e_a, man) \wedge \textit{man}(e_n) \wedge e_n = x \wedge x = vincent]$ *Vincent is cruel*

(214) $\not\models \exists x \exists y \wedge \exists e_a \exists e_n.[\textit{cruelty}(e_a) \wedge \textit{theme}(e_a, x) \wedge \textit{role\_as}(e_a, friend) \wedge \textit{man}(e_n) \wedge e_n = x \wedge x = vincent]$ *Vincent is a cruel friend*

The predicate `role_as` permits to control the inheritance relations which hold for the properties denoted by this type of adjectives and namely that such properties are inherited by the hyponyms of the noun they modify but not by its hyperonyms. This can be expressed with axioms of the form:

(215)

$$\forall N_i s.t. N_i \in Hypo(N) : \texttt{role\_as}(e_a,N) \rightarrow \texttt{role\_as}(e_a,N_i)$$
$$\forall N_i s.t. N_i \in Hyper(N) : \texttt{role\_as}(e_a,N) \not\rightarrow \texttt{role\_as}(e_a,N_i)$$

with $Hypo(N)$ being the set of hyponyms of the modified noun N and $Hyper(N)$ the set of its hyperonyms.
So for example, that a *cruel/unjust leader* is not necessary a *cruel/unjust person* but that a *cruel/unjust leader* who is a *senator* is also a *cruel/unjust senator* is captured through the axiom is (216).

(216)

$$\texttt{role\_as}(e_a,leader) \rightarrow \texttt{role\_as}(e_a,senator)$$
$$\texttt{role\_as}(e_a,leader) \not\rightarrow \texttt{role\_as}(e_a,person)$$

Relational Subsectives

Relational subsective adjectives such as *gastronomical, mechanical, wooden* are all related to nouns. Their semantics can be captured by the representation in (217).

(217) $\lambda N' \lambda Pol \lambda e \lambda x. \exists e_a \exists e_n [\textsf{N}_a(e_a) \wedge Pol(\textit{Rel}(e_a, e_n)) \wedge N'(\lambda S.S)(e_n)(x)]$

Levi (1978) has done a very interesting analysis of complex nominals (CN), i.e. noun-noun compounds and relational adjective-noun combinations. The result of this analysis is that the meaning of complex nominals can be captured by considering two syntactic processes:

- predicate nominalisation
- predicate deletion

According to Levi, the CNs which originated by predicate deletion can be described by 9 semantic relations and those which originated by predicate nominalisation by 4 semantic relations (see Figure 8.4).
Therefore, Levi's relations can be used to further divide this class into subclasses grouping together adjectives which belong to the same ontological class so to be able to specify the decompositional semantics of relational subsective adjectives by defining the relation Rel (e.g. Levi classes can used to group together adjectives for which the Rel relation can be captured by the relation $about$ or $make$ etc.).
As an example, in (218) I show the semantics of the adjectives *gastronomical*, *civil* and

Predicate Deletion		
Predicate	N is Subj(Pred)	Adj is Subj(Pred)
CAUSE	malarial mosquitoes mortal blow	viral infection thermal stress
MAKE	musical clock sebaceous glands	molecular chains stellar configurations
HAVE	musical comedy industrial area	feminine intuition reptilian scales
USE	manual labour solar generator	
BE	mammalian vertebrates professional friends	
IN	marine life autumnal rains	
FOR	avian sanctuary	
FROM	rural visitors solar energy	
ABOUT	criminal policy linguistic lecture	
Predicate Nominalisation		
AGENT	urban planner electrical conductor	
PATIENT		royal gift presidential appointees
PRODUCT	oceanic studies chromatic analyses	clerical error faculty decisions
ACT	musical criticism birth control	parental refusal judicial betrayal

Figure 8.4: Levi's Classification of CNs

solar.

(218)

gastronomical:	$\lambda N'\lambda Pol\lambda e\lambda x\exists e_a\exists e_n.[N_a(e_a)\wedge Pol(\textsf{about}(e_a,e_n))\wedge N'(\lambda S.S)(e_n)(x)]$
provincial:	$\lambda N'\lambda Pol\lambda e\lambda x\exists e_a\exists e_n.[N_a(e_a)\wedge Pol(\textsf{from}(e_a,e_n))\wedge N'(\lambda S.S)(e_n)(x)]$
solar:	$\lambda N'\lambda Pol\lambda e\lambda x\exists e_a\exists e_n.[N_a(e_a)\wedge Pol(\textsf{use}(e_a,e_n))\wedge N'(\lambda S.S)(e_n)(x)]$

Thus the representation of *Something is a gastronomical book*, i.e.

(219) $\exists x\exists e_a\exists e_n.[\textsf{gastronomy}(e_a)\wedge\textsf{about}(e_a,x)\wedge\textsf{book}(e_n)\wedge e_n=x]$

correctly entails that there is an entity x which is a book and is about gastronomy but not that x is gastronomical.

(219)	$\models$	$\exists x.[\textsf{book}(x)]$	*x is a book*
(219)	$\models$	$\exists x\exists e_a.[\textsf{about}(x,e_a)\wedge\textsf{gastronomy}(e_a)]$	*x is about gastronomy*
(219)	$\not\models$	$\exists x[\textsf{book}(x)\wedge\textsf{gastronomy}(x)]$	*x is a book and a gastronomy*
(219)	$\not\models$	$\exists x[\textsf{gastronomical}(x)]$	*x is gastronomical*

8.2.3 Privative Adjectives

Privative adjectives entail that the entity described by the NP is not N, e.g. a fake gun is not a gun. As we saw in Chapter 7, privatives can be subclassified in 7 subclasses by considering their model theoretic semantic, lexical semantic, syntactic and morphoderivational properties. Figure 8.5 shows the different instantiations of the general schema for each of the 7 subsclasses of privatives.

Existence Privatives: Class PRpa1

These adjectives are privative and absolute. Thus, in this case, it is the entity introduced by the adjective that is being quantified over, hence e_a is identified with x (i.e. R_1 is the identity relation). Further, the adjectives in this class negate the existence in the real world of the individual denoted by the modified noun, e.g. a fictitious address is an address which does not exists. This is accounted for by providing the appropriate instantiation for A', e.g. $A' = A(e_a)\wedge\neg\textsf{Exist}(e_a)$.
As an example, I give in (220) the semantic representation assigned to the adjective *fictitious*.

(220) **fictitious:** $\lambda N'\lambda Pol\lambda e\lambda x\exists e_a\exists e_n.[Pol(\textsf{fictitious}(e_a))\wedge Pol(\neg\textsf{Exist}(e_a))\wedge e_a=x$
$\wedge N'(\lambda S.S)(e_n)(x)]$

(221) a. This is a fictitious address
$\exists x\exists e_a\exists e_n.[\textsf{fictitious}(e_a)\wedge\neg\textsf{Exist}(e_a)\wedge e_a=x\wedge\textsf{address}(e_n)\wedge e_n=x])$

b. This address is not fictitious
$\exists x\exists e_a\exists e_n.[\neg\textsf{fictitious}(e_a)\wedge\neg\neg\textsf{Exist}(e_a)\wedge e_a=x\wedge\textsf{address}(e_n)\wedge e_n=x])$

Privative Adjectives	
PRpa1	$\exists e_a \exists e_n [Pol(\mathsf{A}(e_a)) \wedge Pol(\neg \mathsf{Exist}(e_a)) \wedge e_a = x \wedge N'(\lambda S.S)(e_n)(x)]$ e.g. *fictitious*
PRpa2	$\exists e_a \exists e_n [Pol(\mathsf{A}(e_a)) \wedge \mathsf{role\_as}(e_a, n) \wedge Pol(N'(\lambda S.\neg S)(e_n)(x))]$ e.g. *fake*
PRpa4	$\exists e_a \exists e_n [Pol(\mathsf{A}(e_a)) \wedge \mathsf{mod}(e_a, n) \wedge N'(\lambda S.\neg S)(e_n)(x)]$ e.g. *probable*
PRa2	$\exists e_a \exists e_n [Pol(\mathsf{A}(e_a)) \wedge \mathsf{mod}(e_a, n) \wedge N'(\lambda S.\neg S)(e_n)(x)]$ e.g. *eventual*
PRpa3	$\exists e_a \exists e_n [\mathsf{V}_a(e_a)) \wedge Pol(\mathsf{patient}(e_a, x)) \wedge \mathsf{state\text{-}of\text{-}affaires}(e_a, n) \wedge N'(\lambda S.\neg S)(e_n)(x)]$ e.g. *foreseen*
PRa1	$\exists e_a \exists e_n \exists e_{n1} [\mathsf{V}_a(e_a)) \wedge Pol(\mathsf{agent}(e_a, x)) \wedge \mathsf{state\text{-}of\text{-}affaires}(e_a, e_{n1}) \wedge N'(\lambda S.S)(e_{n1})(e_{n1}) \wedge N'(\lambda S.\neg S)(e_n)(x)]$ e.g. *pretended*
PRa3	$Pol(\exists e_a \exists e_n [\mathsf{A}(e_a) \wedge \mathsf{time}(e_a, n) \wedge N'(\lambda S.S)(e_n)(x)])$ e.g. *former*

Figure 8.5: Semantics of Privative Adjectives

Thus the semantic representations assigned to the sentences (221) correctly entail that a *fictitious address* is fictitious and thus does not exist, but that an address that is not fictitious is a (real) address, an address that exists, i.e.

(221a)	$\models \exists x \exists e_a.[\textsf{fictitious}(e_a) \wedge x = e_a]$	*x is fictitious*
(221a)	$\models \exists x \exists e_n.[\neg \textsf{Exist}(x) \wedge \textsf{address}(e_n) \wedge e_n = x]$	*this address does not exist*
(221b)	$\models \exists x \exists e_a.[\neg \textsf{fictitious}(e_a) \wedge x = e_a]$	*x is not fictitious*
(221b)	$\models \exists x \exists e_n.[\textsf{Exist}(x) \wedge \textsf{address}(e_n) \wedge e_n = x]$	*this address exists*

Common Privatives: Class PRpa2

For such adjectives, which are privative and nonabsolute, it is the N property to be denied. This is accounted for by applying the modified noun N to the right polarity (i.e. $Pol_n = \lambda S.\neg S$). Further the fact that these adjectives are nonabsolute, e.g. a fake gun is fake as a gun but not as an object, is captured by instantiating R_2 with the relation $\textsf{role\_as}(e_a, N)$ which provide a link between the property described by the adjective and that described by the modified noun, so that e.g. *fake* is a property of the noun *gun* but not of the individual denoted by it.
For instance, the adjective *fake* is assigned the semantic representation in (222)

(222) **fake:** $\lambda N' \lambda Pol \lambda x \exists e_a \exists e_n.[Pol(\textsf{fake}(e_a)) \wedge \textsf{role\_as}(e_a, n) \wedge Pol(N'(\lambda S.\neg S)(e_n)(x))]$

(223) a. This is a fake gun
$\exists x \exists e_a \exists e_n.[\textsf{fake}(e_a) \wedge \textsf{role\_as}(e_a, gun) \wedge \neg \textsf{gun}(e_n) \wedge e_n = x])$

b. This gun is not fake
$\exists x \exists e_a \exists e_n.[\neg \textsf{fake}(e_a) \wedge \textsf{role\_as}(e_a, gun) \wedge \textsf{gun}(e_n) \wedge e_n = x])$

The representations in (223) allow the following inference patterns to be predicted:

(223a)	$\models \exists x \exists e_a.[\textsf{fake}(e_a) \wedge \textsf{role\_as}(e_a, x)]$	*x is fake*
(223a)	$\models \exists x \exists e_n.[\neg \textsf{gun}(e_n) \wedge e_n = x]$	*x is not a gun*
(223b)	$\models \exists x \exists e_a.[\neg \textsf{fake}(e_a) \wedge x = e_a]$	*x is not fake*
(223b)	$\models \exists x \exists e_n.[\textsf{gun}(e_n) \wedge e_n = x]$	*x is a gun*

Further, R_2 permits to access lexical semantic information, i.e. lexical axioms, which control the of the adjective.

Deverbal Privatives: Class PRpa3 and Class PRa1

The classes PRpa3 and PRa1 includes privatives which are morphologically related to verbs. The class PRpa3 includes members such as *foreseen, predicted* which describe a situation in which the modified noun is the patient of some predictions, while the class PRa1 includes members such as *pretended, seeming* in which the individual denoted by the modified noun is the agent of a pretended event. These difference is accounted for by means of different theta roles. As an example, (224) shows the representation assigned to the adjective *pretended*.

(224) **pretended:** $\lambda N' \lambda Pol \lambda e \lambda x \exists e_a \exists e_n \exists e_{n1} [\textsf{pretend}(e_a) \wedge \textsf{agent}(e_a, x) \wedge \textsf{state-of-affaires}(e_a, e_{n1}) \wedge N'(\lambda S.S)(e_{n1})(e_{n1}) \wedge N'(\lambda S.\neg S)(e_n)(x)]$

The representation of *John is a pretended doctor*, i.e.

(225) $\exists x \exists e_a \exists e_n \exists e_{n1}.[\textsf{pretend}(e_a) \wedge \textsf{agent}(e_a, x) \wedge \textsf{state-of-affaires}(e_a, e_{n1}) \wedge \textsf{doctor}(e_{n1}) \wedge \neg \textsf{doctor}(e_n) \wedge e_n = x])$

correctly entails

(225) $\models \exists x \exists e_a \exists e_n.[\neg \textsf{doctor}(e_n) \wedge e_n = x \wedge x = john]$ *John is not a doctor*

Time Privatives: Class PRa3

Members of this subclass are adjectives such as *former* and *future*. In this case the modified noun is subject to a modality introduced by the adjective, which in this case is accounted for by assigning to the relation R_2 the value $time$.
The semantic representation common to the adjectives of this class is given in (226) for the adjective *former*.

(226) **former:** $\lambda N' \lambda Pol \lambda e \lambda x \exists e_a \exists e_n.Pol([\textsf{former}(e_a) \wedge \textsf{time}(e_a, n) \wedge N'(\lambda S.S)(e_n)(x)])$

(227) a. John is the former president

b. $\exists x \exists e_a \exists e_n [\textsf{former}(e_a) \wedge \textsf{time}(e_a, president) \wedge \textsf{president}(e_n) \wedge e_n = x \wedge x = \textsf{john}]$

Thus, the representation in (227b) allows the following inference patterns to be predicted for the sentence *John is the former president*, i.e.

(227b) $\not\models \exists x \exists e_a \exists e_n.[\textsf{president}(e_n) \wedge \textsf{time}(e_a, president) \wedge \textsf{present}(e_a) \wedge x = \textsf{john}]$ *John is the president*

(227b) $\not\models \exists e_a.[\textsf{former}(e_a) \wedge \textsf{time}(e_a, john)]$ *John is former*

The representation I propose for this subclass of privatives presupposes that each sentence in which such adjectives do not occur has a default value for time. Thus, for instance the inference pattern in (228)

(228) *John is a former president.* $\not\models$ *John is the president.*

can only be accounted for if the base form is assigned the following default representation:

(229) a. John is the president

b. $\exists e_a \exists x [\textsf{president}(x) \wedge \textsf{time}(e_a, president) \wedge \textsf{present}(e_a) \wedge x = \textsf{john}]$

Modal Privatives: Classes PRpa4 and PRa2

Finally, (230) shows an example representation of modal privatives.

(230) **probable:** $\lambda N' \lambda Pol \lambda e \lambda x \exists e_a \exists e_n.[Pol(\textit{probable}(e_a)) \wedge \textit{mod}(e_a, n) \wedge \neg N'(\lambda S.S)(e_n)(x)]$

(231) a. John is a probable candidate
$\exists x \exists e_a \exists e_n [\textit{probable}(e_a) \wedge \textit{mod}(e_a, candidate) \wedge \neg \textit{candidate}(e_n) \wedge e_n = x \wedge x = \textit{john}]$

b. John is an improbable candidate
$\exists x \exists e_a \exists e_n [\neg \textit{probable}(e_a) \wedge \textit{mod}(e_a, candidate) \wedge \neg \textit{candidate}(e_n) \wedge e_n = x \wedge x = \textit{john}]$

The semantic representation of the sentence *John is a probable candidate* correctly predicts that *John is not a candidate* and does not entail *John is probable*, i.e.

(231a)	$\models \exists x \exists e_n.[\neg \textit{candidate}(e_n) \wedge e_n = x \wedge x = \textit{john}]$	*John is not a candidate*
(231a)	$\not\models \exists e_a.[\textit{probable}(e_a) \wedge \textit{mod}(e_a, john)]$	*John is probable*
(231b)	$\models \exists x \exists e_n.[\neg \textit{candidate}(e_n) \wedge e_n = x \wedge x = \textit{john}]$	*John is not a candidate*
(231b)	$\not\models \exists e_a.[\neg \textit{probable}(e_a) \wedge \textit{mod}(e_a, john)]$	*John is improbable*

8.2.4 Plain Nonsubsective Adjectives

Finally, plain nonsubsective adjectives fail to make any prediction about the truth value of the property expressed by the modified noun. Thus for instance, if *John is an alleged murderer*, John might or might not be the murderer.

Plain Nonsubsective Adjectives	
PlNS	$[\textit{say}(e_a) \wedge \textit{patient}(e_a, x) \wedge \textit{topic}(e_a, e_n) \wedge N'(\lambda S.S)(e_n)(e_n)]$ e.g. *alleged*

Figure 8.6: Semantics of Plain Nonsubsective Adjectives

To account for this fact, the semantics of plain nonsubsective adjectives simply specifies that an alleged murderer is an individual x which exists in the universe of discourse and which is alleged to be a murderer. In this case the noun property denotes an eventuality e_n which is not equated with x. R_1 is the identity relation between x and e_a and R_2 is the relation introduced by the adjective $A'(e_a, e_n)$.

(232) **alleged** $\lambda N' \lambda Pol \lambda e \lambda x \exists e_a \exists e_n.[Pol(A'(e_a, e_n)) \wedge x = e_a \wedge N'(\lambda S.S)(e_n)(e_n)]$

Thus the representation assigned to *John is an alleged murderer*,

(233) $\exists x \exists e_a \exists e_n.[\textit{alleged}(e_a, e_n) \wedge x = e_a \wedge \textit{murderer}(e_n) \wedge e_n = e_n \wedge x = \textit{john}])$

correctly predicts the expected inferences, i.e.

(233) $\not\models \lambda x \exists e_n.[\textit{murderer}(e_n) \wedge e_n = x \wedge x = \textit{john}])$ *John is a murderer*
(233) $\not\models \lambda x \exists e_n.[\neg \textit{murderer}(e_n) \wedge e_n = x \wedge x = \textit{john}])$ *John is not a murderer*

The meaning of plain nonsubsective adjectives can be further decomposed by relying on the semantics of the morphologically related Say-verbs. Generally in the analysis of such verbs the whole SC argument is taken to fill the `topic` role of the verb, I propose here another interpretation that better predicts the observed inferential behaviour of such adjectives. My idea is to let the subject of the SC complement fill the `patient` role of the Say-verb and the rest of the SC fill the `topic` role, in this way it is the truth value of the topic which is questioned not the existence of the patient.
Thus for instance, the decompositional representation of *Someone claims John is an infanticide*,

(234) $\exists x \exists y \exists e_a \exists e_n.[\textit{say}(e_a) \wedge \textit{agent}(e_a, y) \wedge \textit{patient}(e_a, x) \wedge \textit{topic}(e_a, e_n) \wedge \textit{infanticide}(e_n) \wedge x = \textit{john}]$

correctly entails that John is an alleged infanticide and, if we assume axioms describing hyponymic relations between noun concepts (e.g. infanticide $\models$ murderer), that John is an alleged murderer but not an alleged terrorist, i.e.

(234) $\models \exists x \exists e_a \exists e_n.[\textit{say}(e_a) \wedge \textit{patient}(e_a, x) \wedge \textit{topic}(e_a, e_n) \wedge \textit{infanticide}(e_n) \wedge x = \textit{john}]$
John is an alleged infanticide
(234) $\not\models \exists x \exists e_n.[\textit{infanticide}(e_n) \wedge e_n = x \wedge x = \textit{john}]$ *John is an infanticide*
(234) $\not\models \exists x \exists e_n.[\neg \textit{infanticide}(e_n) \wedge e_n = x \wedge x = \textit{john}]$ *John is not an infanticide*
(234) $\models \exists x \exists e_a \exists e_n.[\textit{say}(e_a) \wedge \textit{patient}(e_a, x) \wedge \textit{topic}(e_a, e_n) \wedge \textit{murderer}(e_n) \wedge x = \textit{john}]$
John is an alleged murderer
(234) $\not\models \exists x \exists e_a \exists e_n.[\textit{say}(e_a) \wedge \textit{patient}(e_a, x) \wedge \textit{topic}(e_a, e_n) \wedge \textit{terrorist}(e_n) \wedge x = \textit{john}]$
John is an alleged terrorist

8.3 Implementation

The semantics of adjectives presented in this chapter was tested using the BB1 (Blackburn and Bos, 2005) computational semantic framework: a symbolic grammar framework implemented in Prolog dealing with first order inference.
The architecture of the system can be sketched as follows: a simple DCG grammar produces first order semantic representations for input sentences. These representations are then sent together with an appropriate reasoning task to the reasoning module which provides an interface to the model builders MACE[39] and PARADOX[40] and to the theorem provers OTTER [41] and BLIKSEM.[42] Finally, the results are compared with the

[39]MACE 2.0 *http://www-unix.mcs.anl.gov/AR/mace/*
[40]PARADOX 0.4.1 *http://www.math.chalmers.se/∼koen/paradox*
[41]OTTER *http://www.mcs.anl.gov/AR/otter*
[42]BLIKSEM *http://www.mpi-sb.mpg.de/∼bliksem*

expected results.

8.3.1 Grammar and Lexicon

I implemented a DCG grammar fragment which integrates the semantics of nouns and adjectives presented in this chapter. The lexicon contains 42 adjectival items that is, one member of each semantic class defined in Chapter 7. This grammar fragment was then used together with the appropriate lexicon to automatically associate with each sentence a representation of its meaning.

The lexicon maps items belonging to different semantic/syntactic classes to the first order semantic representation shown in Figure 8.1. In particular the lexicon entry for adjectival items contains semantic (e.g. semantic class), syntactic and morphoderivational (e.g. related nouns and/or verbs) information.

```
lexEntry(adj,[symbol:sleep,syntax:[asleep], relVerb:[sleep],
         class:ipa2]).
lexEntry(adj,[symbol:dry,syntax:[dry], relNoun:[wetness],
         class:ipa1]).
lexEntry(adj,[symbol:wet,syntax:[wet], relNoun:[wetness],
         class:ipa1]).
lexEntry(adj,[symbol:big,syntax:[big], relNoun:[size],
         class:spa2]).
lexEntry(adj,[symbol:gastronomical,syntax:[gastronomical],
         relNoun:[gastronomy], class:sa3]).
```

Figure 8.7: Lexicon

This representation supports the construction of the first order semantic representation correspondent to different semantic classes. For instance, the adjective *red* is assigned the semantic representation common for all adjectives of the Ipa2 class instantiated with the value of the `symbol` feature, i.e. *red*, and the adjective *big* is assigned the semantic representation common for all adjectives of the Spa2 class instantiated with the value of the `relNoun` feature, i.e. *size*.
Synonymy is captured by mapping the synonyms to the same semantic representation defined by the value of `symbol` and `class` features.

8.3.2 Lexical Knowledge

Lexical Knowledge pertaining to each lexical items and in particular to each class of adjectives is captured through a set of axioms describing the specific lexical relations adjectives are involved in.

Hyponymy (for example big/giant vs. small/minuscule) is captured by introducing axioms such as:

$$\forall e[\mathrm{Adj}_1(e) \rightarrow \mathrm{Adj}_2(e)]$$

where $\mathtt{Adj}_1$ is an hyponym of $\mathtt{Adj}_2$.

Antonymy is captured by introducing different axioms depending on the type of opposition relation in which the adjectives are involved, i.e. binary, contrary or multiple. The axiom below for example introduces a binary relation between the adjective Adj_1 and its antonym Adj_2:

$$\forall e[\mathrm{Adj}_1(e) \leftrightarrow \neg\, \mathrm{Adj}_2(e)]$$

8.3.3 Reasoning Task

The problem of checking entailment between two sentences can be defined in the following way:

> Given two sentences $S1$ and $S2$ and their logic representation $\phi(S1)$ and $\phi(S2)$ then the first sentence entails the second $S1 \models S2$ iff the formula (235) is valid.

(235) $\phi(S1) \cup LexicalK \implies \phi(S2)$

$LexicalK$ corresponds to the set of axioms which describe the lexical semantics of the lexical items contained in sentence $S1$.

Now, as it is well known, first order logic is undecidable and only a partial solution exists to the problem of checking the validity of a formula. However, as some authors (Blackburn and Bos, 2005) have shown, by interleaving model building with theorem proving the search for a possible solution is facilitated.
Theorem provers in fact can check the validity of a formula directly, i.e. they can prove that a formula is valid but the opposite task, testing that a formula is not valid, is intractable. On the other hand, model builders can test the satisfiability of a logic formula by building a model. Thus model building provides a test for invalidity, if a formula is not valid then there exists a model in which its negation is satisfiable.

In our case, the validity of the entailment between two sentences is tested by sending the appropriate reasoning tasks to model builder and theorem prover, i.e. the formula (236) is sent to the theorem prover for being tested on validity and the formula (237) is sent to the model builder for being tested on satisfiability.

(236) $\phi(S1) \cup LexicalK \implies \phi(S2)$

(237) $\neg(\phi(S1) \cup LexicalK \implies \phi(S2))$

The theorem provers and model builders provided by the BB1 system can be set so that an answer is always returned. This is done by defining a limit to the search time of theorem provers (e.g. 20 seconds) and by setting the maximal size of the domain of model builders (e.g. 30 objects).

Within this framework I have tested whether entailment holds between the sentence pairs in the examples presented in this chapter (about 150 pairs) and compared the results with the expected result. A first evaluation shows that the methodology proposed is promising: the framework could correctly predict all the inferential patterns with model theoretic source and those originated by antonymy. The results for other patterns, in particular those with morphoderivational source, depend on the amount of lexical information implemented (i.e. complexity of syntax considered in the grammar and lexical axioms) which for the moment is very limited.

8.4 Summary

In this chapter I developed a unified first order representation of adjectival lexical semantics. I have shown that the classification outlined in Chapter 7 allows to assign a finer grained semantic to each adjectival item so that inferential patterns with different source can be predicted.

My aim here was not to give an exhaustive decompositional semantics for each adjectival class, but rather to show that each class can be assigned a first order flat representation by instantiating the general scheme for adjectival semantics.

In the next chapter I present the test suite for adjectival inference I have developed based on the results of the previous chapters.

A Test Suite for Adjectival Inference

The detailed analysis of adjectives presented in this thesis has also lead to the construction of a test suite specifically focusing on the interactions between derivational morphology, lexical and compositional semantics and of their impact on the entailment patterns licensed by sentences containing adjectives or their related nouns/verbs.
With the design of this test suite my aim was to create a resource supporting the evaluation of computational systems handling natural language inference and in particular to provide a benchmark against which to evaluate and compare existing semantic analysers.

In this chapter, I present the test suite for adjectival inference I have developed and give some details of its realisation.

9.1 Motivation

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. I 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.

The collections of test data appeared till now, such as the Paraphrase Research Corpus (C. et al., 2004; Dolan et al., 2004), the RTE1 (Dagan et al., 2006), RTE2 (Bar-Haim et al., 2006) data sets and the Text Retrieval Conference TREC (2007) data, created to evaluate systems for wide coverage and robustness, do not satisfy, in my 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.

```
Sentence ID   String   Author  URL  Agency  Date  Web Date

702876  Amrozi accused his brother, whom he called "the wit-
        ness", of deliberately distorting his evidence.
        Darren Goodsir  www.theage.com.au  *   June 5 2003
        2003/06/04

702977  Referring to him as only "the witness", Amrozi
        accused his brother of deliberately distorting his
        evidence.
        Darren Goodsir  www.smh.com.au  Sydney Morning Herald
        June 5 2003    2003/06/04
2108705 Yucaipa owned Dominick's before selling the chain to
        Safeway in 1998 for $2.5 billion.   MICHAEL GIBBS
        www.nwherald.com        *       *      2003/08/23
```

Figure 9.1: IBM paraphrase corpus

```
<pair id="8" entailment="NO" task="IE">
<t>Mangla was summoned after Madhumita's sister Nidhi Shukla,
   who was the first witness in the case.</t>
<h>Shukla is related to Mangla.</h>
</pair>
<pair id="15" entailment="YES" task="IE">
<t>A mercenary group faithful to the warmongering policy of
   former Somozist colonel Enrique Bermudez attacked an IFA
   truck belonging to the interior ministry at 0900 on 26
   March in El Jicote, wounded and killed an interior
   ministry worker and wounded five others.</t>
<h>An interior ministry worker was killed by a mercenary
   group.</h>
</pair>
```

Figure 9.2: RTE-challenge entailment corpus

On the contrary, I believe 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. As the TSNLP (Oepen and Netter, 1995) 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.

In this thesis building on (Amoia and Gardent, 2008), I have addressed the entailment problem from a linguistically based perspective and I have developed a test suite which focuses on a specific linguistic task, namely adjectival inference and addresses this issue deeply. Moreover, I use a well defined notion of entailment as I aim at providing a resource for also evaluating deep semantic analysers based on symbolic methods. Thus, the test suite presented in this chapter 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 in Chapter 7.
In collecting the test data, I 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 I first define the notion of entailment I presuppose, then I describe linguistic task, lexical coverage and realisation of the test suite.

9.2 The Entailment Recognition Task

The idea behind the construction of this test suite is to illustrate the semantic and syntactic behaviour 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 recognised 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 I use corresponds to the notion of logic entailment between the representations of the two sentences:

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

where $\Phi(\mathcal{S}_i)$ corresponds to the logic representation of the sentence S_i.

9.3 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, I build on the results of the previous chapters, in which I 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, I have first individuated a set of general properties of adjectives by merging together, as we have seen in Chapter 7, syntactic properties of adjectives with lexical semantic, model theoretic and morphoderivational properties. Then, I have extracted the inferential patterns which originate from these properties, thus obtaining a set of about 40 inferential patterns. In the following, I describe in detail the patterns annotated in the test suite.

Syntactic Patterns

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

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

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

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

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

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

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

Furthermore, I consider adjectival constructions with clausal complement (SC) such as object embedding, subject embedding, easy/tough constructions.

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

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

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

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

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

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

P7: To-Construction
`N is A SC` ↔ `SC is A to N`

(244) *Ann was clear that the helmet he was wearing saved his life* $\leftrightarrow$ *That the helmet he was wearing saved his life was clear to Ann*

P8: Easy-Construction I
`N is A SC` $\leftrightarrow$ `It is A SC`

(245) *John is easy to talk to* $\leftrightarrow$ *It is easy to talk to John*

P9: Easy-Construction II
`N is A for-PP SC` $\leftrightarrow$ `It is A for-PP SC`

(246) *John is easy for Mary to talk to* $\leftrightarrow$ *It is easy for Mary to talk to John*

Lexical Semantics

The different behaviour shown by adjectives with respect to their antonyms (Cruse, 1986) originates the two entailment relations below:

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

(247) *The dishcloth is wet* $\rightarrow$ *The dishcloth is not dry*

(248) *The mouse is big* $\rightarrow$ *The mouse is not small*

(249) *The table is red* $\rightarrow$ *The mouse is not blue/yellow/green*

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

(250) *The dishcloth is not wet* $\rightarrow$ *The dishcloth is dry*

(251) *The mouse is not small* $\not\rightarrow$ *The mouse is big*

(252) *The table is not red* $\not\rightarrow$ *The table is blue/yellow/green*

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

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

(253) *He is minuscule* $\rightarrow$ *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. Similar patterns have been applied to gradable adjectives to individuate logical polarity.

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

(254) *X is a red table* $\rightarrow$ *X is a red object*

(255) *John is a civil lawyer* $\not\rightarrow$ *John is a civil man*

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

(257) *He is a fictitious hero* $\rightarrow$ *He is a fictitious person*

(258) *John is the alleged strangler* $\rightarrow$ *John is the alleged murderer*

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

(259) *John is a dangerous man and is a husband* $\rightarrow$ *John is a dangerous husband*

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

(260) *It is dangerous to drive in Rome* $\rightarrow$ *It is dangerous to drive fast in Rome*

(261) *It is safe to drive in Rome* $\not\rightarrow$ *It is safe to drive fast in Rome*

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

(262) *It is dangerous to drive fast in Rome* $\not\rightarrow$ *It is dangerous to drive in Rome*

(263) *It is safe to drive fast in Rome* $\rightarrow$ *It is safe to drive in Rome*

Derivational Morphology

Building on Vendler (1963, 1968), Quirk et al. (1985) and other authors, I have collected entailment patterns which have derivational morphology as source. I 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 a verb V which is morphologically related to the adjective A,
ADVa an adverb ADV which is morphologically related to the adjective A,
Nv a noun N which is morphologically related to the verb V.

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

P17: Adjective-Verb Alternation I
`N is Av ↔ N V`

(264) *John is asleep* ↔ *John sleeps*

P18: Adjective-Verb Alternation II
`N is Av ↔ It is possible to V N`

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

P19: Adjective-Verb Alternation III
`N is Av Prep N1 ↔ N V N1`

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

P20: Adjective-Verb Alternation IV
`N1 Va N → N is A`

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

P21: Adjective-Verb Alternation V
`N1 is An2 Nv ↔ N1 V N2`

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

P22: Adjective-Verb Alternation VI
`N is Adj N1 ↔ N1 V_imp ADV`

(269) *This is a fast car* ↔ *This car runs fast*

(270) *This is an expensive review* ↔ *This review costs much*

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

P23: Adjective-ThetaRole_Noun Alternation
`N is An1 ↔ N is N1`

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

P24: Adjective-Event_Noun Alternation
`N is ADVa2 An1 ↔ N’s N1 is A2`

(272) *John is deeply asleep* ↔ *John’s sleep is deep*

P25: Adjective-NonEvent_Noun Alternation
`N is An1 ↔ N’s N1`

(273) *John is polite* ↔ *John’s politeness*

P26: Adjective-CategorialNoun *is*-Alternation
`This is a An1 N` $\leftrightarrow$ `This N is a N1`

(274) *This is an Italian lawyer* $\leftrightarrow$ *This lawyer is an Italian*

P27: Adjective-CategorialNoun *has*-Alternation
`This is a An1 N` $\rightarrow$ `This N has a N1`

(275) *This is a long train* $\leftrightarrow$ *This train has a certain length*

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
`This is An1 N` $\leftrightarrow$ `This N is Rel N1`

(276) *This is a gastronomical dictionary* $\leftrightarrow$ *This is a dictionary about gastronomy*

(277) *They are rural visitors* $\leftrightarrow$ *They are visitors from the country*

(278) *This is a wooden table* $\leftrightarrow$ *This table is made of wood*

P29: Adjective-Adverb Alternation
`N1's Nv is Aadv` $\leftrightarrow$ `N1 V ADV`

(279) *John's smile was cruel* $\leftrightarrow$ *John smiled cruelly*

Vendler (1968) individuates constructions in which the modified noun is substituted with a verbal phrase containing the verb implied by the noun.

P30: `N is A Nv` $\rightarrow$ `N is A to V-imp`

(280) *It is a good meal* $\rightarrow$ *The meal is good to eat*

Model Theoretic Semantics

For model theoretic properties of adjectives, I rely on Kamp and Partee (1995) and Keenan (1987), obtaining the following patterns:

P31: Absoluteness I
`This is A N` $\rightarrow$ `This is A`

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

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

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

P32: Absoluteness II
`X is A SC` $\rightarrow$ `X is A`

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

P33: Absoluteness III: Context-Dependent Properties
`This is NP`$_{meas}$ `A N` $\rightarrow$ `This is NP`$_{meas}$ `A`

(285) *This is a 2 m long line* $\rightarrow$ *This is 2 m long*

P34: Subsectivity I
`This is A N` $\rightarrow$ `This is N`

(286) *This is a red table* $\rightarrow$ *This is a table*

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

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

P35: Subsectivity II
`This is A for N` $\rightarrow$ `This is N`

(289) *Bixi is big for a mouse* $\rightarrow$ *Bixi is a mouse*

(290) *John is anxious for the result* $\not\rightarrow$ *John is a result*

P36: Subsectivity III
`This is A as N` $\rightarrow$ `This is N`

(291) *John is polite as a guest* $\rightarrow$ *John is a guest*

(292) *Bixi is fast as a turtle* $\not\rightarrow$ *Bixi is a turtle*

P37: Privativity
`This is A N` $\rightarrow$ `This is` $\neg$`N`

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

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

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

9.4 Realisation

The test suite 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 partly taken from the literature on adjectives, others are hand coded or were collected from the Web and subsequently simplified, if necessary. The example sentences have been created by generating for each adjectival item[43] sentences representing all inferential patterns possible for that adjective. An attempt was also made to consider an equal number of positive and negative cases of entailment. Figure 9.4 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 judgement about the truth of the entailment between the sentences in the pair. The attribute `entailment` has the values TRUE and FALSE, to respectively tag true and false entailment between the sentences S1 and S2 and TRUE-P to signalise bidirectional entailment, i.e. 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 chapter. So for example, privative patterns are annotated with `inferencePattern=P37`.

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, I use the semantic classification of adjectives presented in Chapter 7.
I 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 judgement of whether the entailment between the sentences in a given pair holds or not. The adjectival class assignment in fact, maps each adjectival item to a semantic representation which is first order and compositional as described in Chapter 8.

9.5 Summary

In this chapter, I presented a test suite specifically created to study the inferential behaviour of English adjectives. I hope it may serve as a resource for the evaluation of systems handling with natural language inference. With the construction of this test suite I want to open the way for the creation of resources which give a deeper insight

[43]The linguistic coverage of the test suite is identical with that of the classification of adjectives presented in Chapter 7 and includes about 500 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="4" 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>
```

Figure 9.3: An example of corpus annotation

in the linguistic phenomena which are responsible for inference. I am aware of the limits of the test items included in this test suite, as I have considered only base cases of entailment. In future work, I want to concentrate on the extension of the test sample by increasing the complexity of the test items to include cases which result from the combination of simpler ones.

Related Work

The work on adjectives presented in this thesis lays the basis for the creation of a general, reusable resource providing domain independent linguistic knowledge for natural language processing and in particular for adjectival inference recognition.

In this chapter, I review some other attempts to encode linguistic knowledge of adjectives in natural language processing systems, namely two computational lexica based on ontological organisation of adjectival meaning, MicroCosmos (Raskin and Nirenburg, 1995) and SIMPLE (Peters and Peters, 2000), and three broad coverage lexical resources, WordNet (Miller, 1998), FrameNet (Baker et al., 1998; Ruppenhofer et al., 2006) and ComLex (Grishman et al., 1994). Thus, I compare them with the approach presented in this thesis.

10.1 MicroCosmos

Raskin and Nirenburg (1995) describe the methodology used to encode adjectival entries in the lexicon of the MicroCosmos semantic analyser (Beale et al., 1995), which is a component of a knowledge-based machine translation system (Nirenburg et al., 1992). The MicroCosmos lexicon contains 6000 entries for English and 1500 entries for Spanish adjectives. The adjectives are organised in an ontology which distinguishes between the following three main adjectival classes:

- `Scalar Adjectives`
  The meaning of scalar adjectives (e.g. *red, large, big*) corresponds to a *property-value* pair, where the *property* is the ontological property described by the adjective and the *value* is a region in the scale[44] which represents the range of that property.
  The meaning of an [A N]$_{AP}$ phrase corresponds to the meaning of the noun, i.e. a frame corresponding to an instance of an ontological concept, in which the slot corresponding to the adjectival property is filled with its value.

[44]MicroCosmos distinguishes between continuous (e.g. length) and discrete (e.g. colour) scales.

Consider, as an example, how the meaning of *big house* is captured.
If the meaning of the adjective *big* corresponds to the property-value pair (P:V) *SIZE:>0.75*, meaning that the adjective points to a region in the length-scale with a value greater than 0.75, then the meaning of the $[\text{A N}]_{AP}$ phrase is represented by HOUSE$\{SIZE :> 0.75\}$, a representation which in Description Logic corresponds to the formula: $HOUSE \sqcap \exists SIZE.BIG$.
Scalar adjectives are further divided in subclasses as shown in Figure 10.1.

SubClass	**Property**	**Examples**
Attitude-based	Evaluation, Salience	*good, superb, important*
Numerical Scale	Size, Weight, Age, ...	*big, heavy, young*
Literal Scale	Colour, Shape, ...	*red, oval*
Member	Set Membership	*authentic, fake*

Figure 10.1: Taxonomy of scalar adjectives in MicroCosmos

- `Denominal Adjectives`
  The meaning of denominal adjectives (e.g. *atomic, civil, gastronomical*) is related to the meaning of the nouns they are derived from through the *pertain-to* relation. Consider, for example, the entry for the adjective *medical*:

  (medical-Adj → (Cat:adj)
  (Sem: (pertain-to medicine))
)

  Thus, the meaning assigned to *medical devise* is

  DEVISE $\sqcap$ PERTAIN-TO.MEDICINE

- `Deverbal Adjectives`
  The meaning of deverbal adjectives (e.g. *eager, abusive, readable*) is related to the meaning of the verb they are derived from. For example, the entry for the adjective *abusive* is as given below:

  (abusive-Adj → (Cat:adj)
  (Sem: abuse
  (agent:Modified_Noun)
)
  (attitude-value:<0.25)
)

The MicroCosmos lexicon was designed to deal with a large set of semantic phenomena and seems promising in coping with adjectival inferences based on morphoderivational knowledge or involving privative adjectives, but fails to generalise to other types of adjectives such as plain nonsubsective. However, in order to be usable, the MicroCosmos lexicon should be linked to the MicroCosmos ontology. To my knowledge, the ontology was designed only for the restricted domain of company merges and acquisitions and consists of 5000 concepts. Thus, it is unclear whether the detailed knowledge about adjectives has ever been fully implemented.

10.2 Simple

The SIMPLE project (SIMPLE, 2000) had the aim of extending with semantic information the set of lexica, built for 12 European languages by the PAROLE Consortium, which provides morphological and syntactic information for about 20,000 lexical entries.

1- `Intensional`			
	1.1- `Temporal`		
		1.1.1- `Past`	*former*
		1.1.2- `Present`	*present*
		1.1.3- `Future`	*future*
	1.2- `Modal`		
		1.2.1- `Epistemic`	*certain*
		1.2.2- `Deontic`	*necessary*
		1.2.3- `Possibility`	*potential*
	1.3- `Emotive`		*poor*
	1.4- `Manner`		*beautiful*
	1.5- `Object-Related`		*criminal*
	1.6- `Emphasizer`		*outright*
2- `Extensional`			
	2.1- `Psychological Property`		*crazy*
	2.2- `Social Property`		*catholic*
	2.3- `Physical Property`		*soft*
	2.4- `Temporal Property`		*sudden*
	2.5- `Intensifying Property`		*heavy*
	2.6- `Relational Property`		*similar*

Figure 10.2: SIMPLE ontology for adjectives

The classification of adjectives in SIMPLE (Peters and Peters, 2000) is based on the ontology depicted in Figure 10.2. A lexical entry for an adjective is characterised by

a set of semantic and syntactic information. Semantic information describes: (i) the hierarchy of ontological properties expressed by the particular adjective, (ii) whether the adjective is intersective or subsective; (iii) whether the adjective has a persistent duration (i.e. is stable) or not.
Syntactic information describes adjectival features such as (i) predicative/attributive usage, and (ii) gradability. Figure 10.3 shows the lexical entry for the adjective *wet* as it is coded in SIMPLE. *Wet* is an adjective describing the static physical property of *Wetness*; this adjective is extensional and intersective and can be used both attributively and predicatively.

SIMPLE has actually added semantic information to approximately 3500 lexical entries (about 10,000 senses) for each of the 12 European languages considered in the project. However, it is not clear how many adjectives the English lexicon covers. Peters (2000) reports that in March 2000, 500 adjectives had not yet been included in the lexicon.

Template Type:	Physical property
Template Supertype:	Extensional
Intersective/subsective:	Intersective
Meaning Components:	Wetness, Plus
Syntactic Type:	Attributive-and-predicative
Gradability:	yes
Duration:	persistent

Figure 10.3: SIMPLE lexical entry for *wet*

10.3 WordNet

As described in Miller (1998), WordNet lists about 16,500 adjectival forms, among these there are also nouns, participles and prepositional phrases which can function as noun modifiers. Adjectives in WordNet are divided into two different categories:

- `descriptive` adjectives, i.e. *red, beautiful, old, possible*
- `relational` adjectives, i.e. those related to nouns such as *stellar, electrical*

Each class is fully individuated by a different set of semantic pointers or relations. The characteristic property which individuates `descriptive` adjectives is that they are related to other adjectives through antonymic relations, which can be direct or indirect, i.e. mediated by clusters built on semantic similarity between adjectives. `Descriptive` adjectives which express values of attributes are further related to the noun they are attributes of by the pointer *is-a-value-of*.

Sense 1	
heavy	(of comparatively great physical weight or density; "a heavy load"; “lead is a heavy metal”; “heavy mahogany furniture”)
	Relations:
	Antonyms $\Longrightarrow$ { light}
	Is a value of $\Longrightarrow$ weight(Sense 1)
	Derivationally related forms:
	Related to noun $\Longrightarrow$ heaviness(Sense 1)

Figure 10.4: WordNet Representation of Descriptive Adjectives: *heavy*

`Relational` adjectives do not have antonyms but are related to nouns through the *Relating-or-Pertaining-to* relation.

Sense 1	
dental(prenominal),	(of or relating to the teeth; "dental floss”)
	Relations:
	Pertains to noun $\Longrightarrow$ tooth(Sense 1)
	Antonyms $\Longrightarrow \emptyset$
	Is a value of $\Longrightarrow \emptyset$

Figure 10.5: WordNet Representation of Relational Adjectives: *dental*

Moreover, WordNet offers some syntactic information, even if very limited, about the attributive/predicative use of an adjective. More precisely, an adjective can be characterised by the following features:

- any: the adjective has both predicative and attributive usages
- predicative: the adjective has only predicative usage, e.g. *awake* (predicate)
- prenominal: the adjective has only attributive usage, e.g. *previous* (prenominal)
- postnominal: the adjective can only be used in postnominal position, e.g. *elect* (postnominal).

WordNet can help predict adjectival inferential patterns based on morphoderivational properties and perhaps some cases of antonymy but cannot be used alone to deal with more complex cases of inference, as semantically finer-grained information is absent. For example, information about gradability of adjectives or on whether they are intersective, subsective, etc., is not implemented.

10.4 FrameNet

FrameNet (Baker et al., 1998; Ruppenhofer et al., 2006) is an on-line lexical database based on frame semantics (Fillmore, 1976). It contains about 6000 fully annotated lexical items described through 800 semantic frames.
Adjectives can also evoke frames: their semantics are related to the semantics of the situation expressed by the frame.

Biological_urge Frame
Definition: An *Experiencer* is in a *State* where a biological urge is signalling the need to perform a certain action such as eating, drinking, stopping exertion, sleeping, regurgitating or having sex. The *Duration*, *Time*, *Place* or *Degree* of intensity associated with the occurrence of the urge may also be indicated. In this frame, the *Experiencer* does not necessarily want to perform the action that would fulfil the urge, and there is no specific focus associated with the urge. **Examples:** Mike$_{Experiencer}$ feels hungry$_{State}$ Sam$_{Experiencer}$ was thirsty$_{State}$ [after dinner]$_{Time}$

Figure 10.6: FrameNet: Biological_urge Frame

Consider, for example, the frame *Biological_urge* described in Figure 10.6 which has the following members:
amorous.a, arousal.n, aroused.a, beat.a, bilious.a, bone-weary.a, bushed.a, dog-tired.a, drowsiness.n, drowsy.a, enervated.a, exhausted.a, exhaustion.n, famished.a, fatigue.n, fatigued.a, horny.a, hunger.n, hungry.a, in the mood.a, knackered.a, nausea.n, nauseated.a, nauseous.a, parched.a, peckish.a, pooped.a, queasiness.n, queasy.a, randy.a, ravenous.a, sick.a, sleepiness.n, sleepy.a, somnolence.n, somnolent.a, soporific.a, thirst.n, thirsty.a, tired.a, tiredness.n, tuckered out.a, tuckered.a, turned on.a, weariness.n, weary.a, worn out.a.
FrameNet associates to each adjectival item a semantic frame but no syntactic information is directly available. FrameNet neither annotates adjectives for their usage, i.e. if the adjectives can be used in predicative and/or attributive position, nor gives any indication on whether they are gradable or not. Moreover, a frame groups semantically similar words together, so that for example, adjectives which are antonyms are listed in the same frame and cannot be distinguished. No further information is available about adjective inferential behaviour, e.g. whether the adjectives are intersective, subsective, etc.

Thus, the information about adjectives present in FrameNet alone cannot be directly used to handle adjectival inference; it should be integrated with other resources such as, for example WordNet, to extract information such as morphological derivations (relational adjectives e.g. *atomic* are not annotated in FrameNet as they are considered as non-frame-bearing elements), antonyms, etc.
Finally, a consideration on the coverage. Most adjectives included in FrameNet (which are about 1000) belong to limited domains, namely to *emotion* and *evaluation*-related frames.

10.5 Comlex

Comlex Syntax[45] (Grishman et al., 1994) is a computational lexicon developed by the Proteus Project at New York University. This lexicon contains detailed syntactic information. It includes 8000 entries for adjectives and contains:

- information about the syntactic frames in which the adjectives participate, i.e. *tough* construction, *it* extraposition, subject/object embedding
- information about the usage of the adjective, i.e. predicative, attributive, postnominal
- information on whether the adjective is gradable or not, i.e. gradable (-er/-est or more/most), not gradable.

However, the Comlex lexicon does not contain any semantic information. Therefore, it cannot directly be used to cope with adjectival inference. However, it can help in classifying adjectives on the basis of their syntactic properties.

10.6 Discussion

The computational classifications of adjectives presented in this chapter are limited in their capability of handling adjectival inference as they are based on a limited set of adjectival features.

Lexical resources such as WordNet, FrameNet and Comlex are wide coverage and support robust processing but implement only one type of knowledge, semantic (e.g. WordNet, Framenet) or syntactic (e.g. ComLex).

On the other hand, computational lexica such as Microcosmos and Simple represent attempts to provide a syntax-semantic interface for NLP systems encoding lexical knowledge about adjectives in ontologies, but they were never fully implemented so that it is difficult to evaluate them.

However, the selectional criteria used in both systems to classify adjectives are model theoretic distinctions such as intensional/extensional, gradability, predicative/attributive

[45] http://nlp.cs.nyu.edu/comlex

patterns and morphological distinctions such as denominal/deverbal. SIMPLE also uses stability/dynamicity criteria. Now, this set of criteria is not as fine grained as the one presented in this thesis. So for example, it seems that neither of the two ontologies, MicroCosmos and Simple, can account for the different kinds of antonymic relations. Also, it is not clear how intensional adjectives are handled. Further, none of these ontologies handles plain nonsubsective adjectives.

On the contrary, the classification of adjectives presented in the previous chapters being based on selectional criteria which integrate semantics, lexical semantics, syntax and derivational morphology provides a more fine grained characterisation of the adjectival classes. Further, each adjectival class is associated with the set of inference patterns that originate from the properties of the members of that class and with a fine grained first order semantic representation which integrates model theoretic, lexical semantic and morphoderivational properties and thus supports the prediction and recognition of all the inference patterns allowed for the class.

10.7 Summary

In this chapter, I gave an overview of existent computational lexical resources for adjectives. In particular, I compared the approach presented in this thesis with ontological-based lexical resources such as MicroCosmos and Simple and with wide coverage lexical resources such as WordNet and FrameNet.

Part III

Conclusions

Conclusions

In this thesis I have investigated how lexical resources based on the organisation of lexical knowledge in classes, which share common (syntactic, semantic, etc.) features, support natural language processing and in particular symbolic recognition of textual entailment.

First I have presented a robust and wide coverage approach to paraphrase recognition based on Levin's (1993) classification of English verbs. Then, I have shown that by extending Levin's framework to general inference patterns, a classification of English adjectives can be obtained that compared with previous approaches, provides a more fine grained semantic characterisation of their inferential properties and thus lays the basis for a computational treatment of adjectival inference.

In this chapter, I review the results of this thesis and conclude with some pointers for future research.

11.1 Verb Paraphrase Recognition

Handling paraphrases is a central issue when developing natural language processing applications such as question-answering, multi-document summarisation and text generation. Thus, paraphrases have recently triggered renewed research interest and much of the research in this field is based on statistical methods (Lin and Pantel, 2001; Barzilay and Lee, 2003; Shinyama et al., 2002; Glickman and Dagan, 2003; Ravichandran and Hovy, 2002) and has concentrated on the (semi)automatic acquisition of paraphrases from corpora.

The approach to paraphrases recognition I have presented in this thesis is quite different, it is linguistic-based and aims at modelling a framework which can assign to two different paraphrastic constructions one and the same semantic representation. My approach focuses on the recognition of verb paraphrases and builds on Levin's classification of English verbs, one of the major attempts in integrating syntax with lexical semantics.

Driven by the idea that similar syntactic properties give a clue to similar semantic behaviour, Levin builds semantic classes of English verbs by focusing on syntactic alternations, i.e. by considering all the different syntactic patterns allowed for verbs that are associated with the same semantic content:

(296) a. *The key opens the safe* ↔ *The safe opens with the key*

b. *I give books to John* ↔ *I give John books*

Verbs in the same verb class share similar semantic properties.

In this thesis I have argued, that the particular organisation of verbal meaning proposed by Levin, and in particular the fact that each verb class and thus each verb is associated with a set of alternations, i.e. paraphrastic patterns possible for the members of the class, is suitable to be used for the automatic acquisition and recognition of paraphrastic patterns of verbs. Therefore, I have shown how the paraphrastic patterns originating from Levin's classification, which is now a resource available on the web (VerbNet (Kipper et al., 2000)), can be extended with linguistic information about synonyms, troponyms, etc. extracted from WordNet (Fellbaum, 1998), and used for the automatic acquisition of paraphrastic patterns of verbs.
Further, I have shown how this information can be linked to a robust parser such as XIP (Xerox Incremental Parser (Ait-Mokhtar et al., 2002)) and can be used to recognise alternation paraphrases. The parser was evaluated against a TestSuite of pairs of paraphrastic sentences yielding the encouraging result that textual equivalence was recognised in 96% of the cases.

The symbolic approach to paraphrasing I presented in this thesis is a first step towards a robust symbolic treatment of paraphrases. In fact, by integrating linguistic information extracted from existing large scale lexical resources such as VerbNet and WordNet with a robust parser like XIP, the proposed approach allows wide coverage extraction and recognition of paraphrastic patterns of verbs.

11.2 Linguistic-Based Recognition of Adjectival Entailment

Building on the result of the first part, I have extended Levin's methodology to classify English adjectives.
Levin (1993) bases her classification of English verbs on a set of 79 alternations describing the paraphrastic patterns in which the verbs can participate. By observing that paraphrases are a subset of the more general set of inference patterns linguistical items can present, I have proposed an important extension to Levin's framework namely, I have argued that linguistic features can be better described by considering the set of inference patterns they display. Further, contrary to Levin's alternations, these patterns are linguistically principled as each type of inference is motivated by the linguistic property which is the source of it.

I have applied this idea to the semantic classification of English adjectives.
First, I have individuated a set of features which are fundamental for the description of adjectival behaviour. This set was obtained by integrating the selectional criteria used in traditional classifications (Huddleston, 1984; Quirk et al., 1985; Vendler, 1963, 1968; Kamp, 1975; Keenan and Faltz, 1985; Keenan, 1987; Chierchia and Connell-Ginet, 1990; Kamp and Partee, 1995) of adjectives, which are mainly concerned with model theoretic semantic and syntactic features, with lexical semantic and morphoderivational properties.
Then, I have collected the set of inference patterns these properties generate. The obtained classification includes 42 classes which are characterised by a different set of inferential patterns. Thus, a semantic class is defined by the set of inferential patterns the items in the class can participate in.

Finally, I have assigned a semantic representation to each class which accounts for model theoretic semantic but also for morphoderivational and lexical semantic properties of the members of the class. As I use first order representations for all adjectival classes, the traditional distinction in intensional and extensional properties is avoided, thus allowing the proposed classification to be efficiently used in natural language processing systems.

Another noteworthy result of this thesis is the test suite I have developed as a resource for the evaluation of NLP applications handling inference (specifically adjectival inference). Contrary to other corpora developed to this end, such as the Paraphrase Research Corpus (C. et al., 2004; Dolan et al., 2004), the RTE1 (Dagan et al., 2006), RTE2 (Bar-Haim et al., 2006) and the Text Retrieval Conference TREC (2007) data sets, this test suite has a clear defined coverage and contains annotations for inferential patterns thus permitting a better control over the analysed data. With the construction of this test suite, my aim was to open the way for the creation of resources which give a deeper insight in the linguistic phenomena which are responsible for inference.

11.3 Future Work

The main contribution of this work is a detailed analysis of the interactions between derivational morphology, lexical and compositional semantics and of their impact on the entailment patterns licensed by sentences containing adjectives or their related nouns or verbs. The approach presented here lays the basis for a computational treatment of adjectival inference in that it provides a fine grained characterisation of the various types of inferential patterns licenced by adjectives. The work on adjectives and the ideas proposed in this thesis open the way to some very interesting future research possibilities. In future work, I believe the following further points are worth investigating.

First, the classification presented here is not comprehensive, the coverage is limited to 500 adjectival items and the set of inferential patterns do not fully cover the set of adjectival properties. For example, adjective PP complementation which is a rich source of inference patterns was not analysed:

(297) a. The composer Salieri was contemporary with Mozart

b. *The composer Salieri and Mozart were contemporaneous*

Thus, I want to examine whether the classification can be further detailed and even finer grained classes identified, thereby permitting the creation of syntactically and semantically more homogeneous adjectival classes by extending the coverage and set of adjectival features and thus the set of inference patterns considered.

Second, I want to apply statistical methods (Lapata, 2001; Briscoe and Carroll, 1997; Brew and im Walde, 2002; Korhonen and Briscoe, 2004) to automatically extend the set of adjectives which present a particular combination of inference patterns in order to implement a wide coverage lexical resource similar to VerbNet supporting entailment recognition for adjectives.

A third point of interest concerns the integration of the compositional semantics proposed in this thesis into a robust semantic processing system. I plan to integrate this semantics into the CCG2Sem semantic parsing system (Bos, 2005) and to investigate in how far, this would help deal with real text entailment recognition.

Further, it would be interesting to see whether the proposed classification can be combined with ontological information, by mapping it to such a lexical semantics ontology as e.g. WordNet, FrameNet, thus providing a general task and domain independent ontology based on linguistic principles in the tradition of the Upper Model Ontology (Bateman, 1990, 1994).

Finally, I want to extend the methodology used here to the semantic classification of other categories such as e.g. nouns and adverbs. It would be also interesting to see how the proposed extension to Levin's methodology, namely to consider inference patterns instead of only paraphrastic patterns, can be applied to verbs so to refine Levin's classes, thus permitting entailment between verbs to be checked.

English Adjectival Classes

A.1 Intersective Adjectives

A.1.1 Class Ipa1

Members

blue1, black1, green1, grey1, orange1, red1, white1, yellow1

Features

Semantic Properties			
Subsectivity	Absoluteness	Transparency	
S	A	T	
Lexical Semantic Properties			
Gradability	Very-Mod	Antonymy	Stability
G	-	M	S
Syntactic Properties			
Pred/Attr Construction	For/As Construction	Complementation	NP_{meas}
PA	-	-	-
Morphoderivational Properties			
N_a			

Inference Patterns

P1: This table is red ↔ This is a red table
P10: This table is red → This table is not blue
P13: This is a red car → This is a red object
P14: This is a red vehicle and is a car → This is a red car
P25: It is clear that this table is red ↔ The redness of this table is clear
P31: This is a red table → This is red
P34: This is a red table → This is a table

Semantics

$\lambda N' \lambda Pol \lambda e \lambda x \exists e_a \exists e_n \exists d[Pol(\mathsf{N}_a(e_a)) \wedge \textit{extent}(e_a, d) \wedge \textit{has-p}(e_a, x) \wedge N'(\lambda S.S)(e_n)(x)]$

(e.g. *red*)

$\lambda N' \lambda Pol \lambda e \lambda x \exists e_a \exists e_n \exists d[Pol(\textit{red}(e_a)) \wedge \textit{extent}(e_a, d) \wedge \textit{has-p}(e_a, x) \wedge N'(\lambda S.S)(e_n)(x)]$

A.1.2 Class Ipa2

Members

dry1, empty1, full1, lighted1

Features

Semantic Properties			
Subsectivity	Absoluteness	Transparency	
S	A	T	
Lexical Semantic Properties			
Gradability	Very-Mod	Antonymy	Stability
G	-	B	D
Syntactic Properties			
Pred/Attr Construction	For/As Construction	Complementation	NP$_{meas}$
PA	-	-	-
Morphoderivational Properties			
N$_a$			

Inference Patterns

P1: This is a dry dishcloth ↔ This dishcloth is dry
P10: This dishcloth is dry → This dishcloth is not wet
P11: This dishcloth is not dry → This dishcloth is wet
P13: This dishcloth is dry → This object is dry
P14: This is a dry object and is a dishcloth → This is a dry dishcloth
P25: It is clear that this dishcloth is dry ↔ The dryness of the dishcloth is clear
P31: This dishcloth is dry → This is dry
P34: This dishcloth is dry → This is a dishcloth

Semantics

$\lambda N' \lambda Pol\ \lambda e \lambda x \exists e_a \exists e_n \exists a \exists d.[\mathsf{N}_a(e_a) \wedge \textsf{\textit{extent}}(e_a, d) \wedge \textsf{\textit{absStandard}}(a, , n_a) \wedge Pol(\textsf{\textit{low/high}}(d, a))$
$\wedge \textsf{\textit{has-p}}(e_a, x) \wedge N'(\lambda S.S)(e_n)(x)]$

(e.g. *dry*)

$\lambda N' \lambda Pol\ \lambda e \lambda x \exists e_a \exists e_n \exists a \exists d.[\textsf{\textit{wetness}}(e_a) \wedge \textsf{\textit{extent}}(e_a, d) \wedge \textsf{\textit{absStandard}}(a, wetness) \wedge Pol(\textsf{\textit{low}}(d, a))$
$\wedge \textsf{\textit{has-p}}(e_a, x) \wedge N'(\lambda S.S)(e_n)(x)]$

A.1.3 Class Ipa3

Members

anterior1, backward1, convergent1, divergent1, forward1, posterior1, straight2, unlighted1, unbreakable1

Features

Semantic Properties					
Subsectivity		Absoluteness		Transparency	
S		A		T	
Lexical Semantic Properties					
Gradability	Very-Mod		Antonymy		Stability
G	-		B		S
Syntactic Properties					
Pred/Attr Construction	For/As Construction		Complementation		NP_{meas}
PA	-		-		-
Morphoderivational Properties					
N_a					

Inference Patterns

P1: This is a convergent sequence ↔ This sequence is convergent
P10: This sequence is convergent → This sequence is not divergent
P11: This sequence is not convergent → This sequence is divergent
P13: This sequence is convergent → This entity is convergent
P14: This is a divergent entity and is a sequence → This is a convergent sequence
P25: It is clear that this sequence is convergent ↔ The convergence of the sequence is clear
P31: This sequence is convergent → This is convergent
P34: This sequence is convergent → This is a sequence

Semantics

$\lambda N' \lambda Pol\, \lambda e \lambda x \exists e_a \exists e_n \exists a \exists d.[\textsf{N}_a(e_a) \wedge \textsf{extent}(e_a, d) \wedge \textsf{absStandard}(a, , n_a) \wedge Pol(\textsf{low/high}(d, a)) \wedge \textsf{has-p}(e_a, x) \wedge N'(\lambda S.S)(e_n)(x)]$

(e.g. *straight*)

$\lambda N' \lambda Pol\, \lambda e \lambda x \exists e_a \exists e_n \exists a \exists d.[\textsf{straightness}(e_a) \wedge \textsf{extent}(e_a, d) \wedge \textsf{absStandard}(a, straightness) \wedge Pol(\textsf{high}(d, a)) \wedge \textsf{has-p}(e_a, x) \wedge N'(\lambda S.S)(e_n)(x)]$

A.1.4 Class Ipa4

Members

adhesive1, alcoholic11, curly1, flat1, fluffy1, fragile1, fragrant1, greasy1, impure1, musical2, nonadhesive1, nonalcoholic1, nonfat1, pointless1, pure1, raw3, rough1, sharp2, smelly1, smooth1, sour2, spicy1, spiny1, sweet1, tasteful1, tasteless1, uneven1, unmusical1, viscid1

Features

Semantic Properties			
Subsectivity	Absoluteness	Transparency	
S	A	T	
Lexical Semantic Properties			
Gradability	Very-Mod	Antonymy	Stability
G	+	B	S
Syntactic Properties			
Pred/Attr Construction	For/As Construction	Complementation	NP_{meas}
PA	-	-	-
Morphoderivational Properties			
N_a			

Inference Patterns

P1: This is a musical boy ↔ This boy is musical
P10: This boy is musical → This boy is not unmusical
P11: This boy is not musical → This boy is unmusical
P13: This boy is musical → This person is musical
P14: This is a musical person and is a boy → This is a musical boy
P25: It is clear that this boy is musical ↔ The musicalness of the boy is clear
P31: This boy is musical → This is musical
P34: This boy is musical → This is a boy

Semantics

$\lambda N' \lambda Pol\ \lambda e \lambda x \exists e_a \exists e_n \exists a \exists d.[\mathsf{N}_a(e_a) \wedge \textit{\textsf{extent}}(e_a, d) \wedge \textit{\textsf{absStandard}}(a, , n_a) \wedge Pol(\textit{\textsf{low/high}}(d, a))$
$\wedge \textit{\textsf{has-p}}(e_a, x) \wedge N'(\lambda S.S)(e_n)(x)]$

(e.g. *spicy*)

$\lambda N' \lambda Pol\ \lambda e \lambda x \exists e_a \exists e_n \exists a \exists d.[\textit{\textsf{spicyness}}(e_a) \wedge \textit{\textsf{extent}}(e_a, d) \wedge \textit{\textsf{absStandard}}(a, spicyness)$
$\wedge Pol(\textit{\textsf{high}}(d, a)) \wedge \textit{\textsf{has-p}}(e_a, x) \wedge N'(\lambda S.S)(e_n)(x)]$

A.1.5 Class Ipa5

Members

bright1, clean1, clear11, cloudy3, dirty1, drunk1, dull2, dusty1, garrulous1, good-natured1, healthy1, hungry1, ill1, irritable1, noisy1, nonslippery1, peaceful1, placid1, quiet2, ripe1, sick1, satiate1, shaky1, slippery1, sober1, stable1, stormy1, sunny1, taciturn1, tame3, testy1, thirsty2, tidy1, unhealthy1, unpeaceful1, unripe1, well1, wet1, wild2

Features

Semantic Properties			
Subsectivity	Absoluteness	Transparency	
S	A	T	
Lexical Semantic Properties			
Gradability	Very-Mod	Antonymy	Stability
G	+	B	D
Syntactic Properties			
Pred/Attr Construction	For/As Construction	Complementation	NP$_{meas}$
PA	-	-	-
Morphoderivational Properties			
N$_a$			

Inference Patterns

P1: This is a clean dishcloth ↔ This dishcloth is clean
P10: This dishcloth is clean → This dishcloth is not dirty
P11: This dishcloth is not clean → This dishcloth is dirty
P13: This dishcloth is clean → This object is clean
P14: This is a clean object and is a dishcloth → This is a clean dishcloth
P25: It is clear that this dishcloth is clean ↔ The cleanness of the dishcloth is clear
P31: This dishcloth is clean → This is clean
P34: This dishcloth is clean → This is a dishcloth

Semantics

$\lambda N' \lambda Pol\ \lambda e \lambda x \exists e_a \exists e_n \exists a \exists d.[\textsf{N}_a(e_a) \wedge \textsf{extent}(e_a, d) \wedge \textsf{absStandard}(a, , n_a) \wedge Pol(\textsf{low/high}(d, a))$
$\wedge \textsf{has-p}(e_a, x) \wedge N'(\lambda S.S)(e_n)(x)]$

(e.g. *bright*)

$\lambda N' \lambda Pol\ \lambda e \lambda x \exists e_a \exists e_n \exists a \exists d.[\textsf{brightness}(e_a) \wedge \textsf{extent}(e_a, d) \wedge \textsf{absStandard}(a, brightness)$
$\wedge Pol(\textsf{high}(d, a)) \wedge \textsf{has-p}(e_a, x) \wedge N'(\lambda S.S)(e_n)(x)]$

A.1.6 Class Ipa6

Members

androgynous1, carnivorous1, female1, gaseous1, herbivorous1, insectivorous1, liquid1, male1, omnivorous1, solid1

Features

Semantic Properties			
Subsectivity	Absoluteness	Transparency	
S	A	T	
Lexical Semantic Properties			
Gradability	Very-Mod	Antonymy	Stability
NG	-	M	S
Syntactic Properties			
Pred/Attr Construction	For/As Construction	Complementation	NP_{meas}
PA	-	-	-
Morphoderivational Properties			
N_{cat}			

Inference Patterns

P1: This animal is carnivorous $\leftrightarrow$ This is a carnivorous animal
P10: This animal is carnivorous $\rightarrow$ This animal is not herbivorous
P13: This is a carnivorous mammal $\rightarrow$ This is a carnivorous animal
P14: This is a carnivorous animal and is a mammal $\rightarrow$ This is a carnivorous mammal
P26: This animal is carnivorous $\leftrightarrow$ This animal is a carnivore
P31: This is a carnivorous animal $\rightarrow$ This is carnivorous
P34: This is a carnivorous animal $\rightarrow$ This is an animal

Semantics

$\lambda N' \lambda Pol \lambda e \lambda x \exists e_a \exists e_n.[Pol(N_{cat}(e_a)) \wedge e_a = x \wedge N'(\lambda S.S)(e_n)(x)]$

(e.g. *carnivorous*)

$\lambda N' \lambda Pol \lambda e \lambda x \exists e_a \exists e_n.[Pol(\mathit{carnivore}(e_a)) \wedge e_a = x \wedge N'(\lambda S.S)(e_n)(x)]$

A.1.7 Class Ipa7

Members

oval1, *quadrate1*, *rectangular1*, *round1*, *triangular1*

Features

Semantic Properties			
Subsectivity	Absoluteness	Transparency	
S	A	T	
Lexical Semantic Properties			
Gradability	Very-Mod	Antonymy	Stability
NG	-	M	S
Syntactic Properties			
Pred/Attr Construction	For/As Construction	Complementation	NP_{meas}
PA	-	-	-
Morphoderivational Properties			
N_a			

Inference Patterns

P1: This is a rectangular table $\leftrightarrow$ This table is rectangular
P10: This table is rectangular $\rightarrow$ This table is not oval
P13: This is a rectangular table $\rightarrow$ This is a rectangular object
P14: This is a rectangular object and is a table $\rightarrow$ This is a rectangular table
P25: It is clear that this table is rectangular $\leftrightarrow$ The rectangularity of the table is clear
P31: This is a rectangular table $\rightarrow$ This is rectangular
P34: This is a rectangular table $\rightarrow$ This is a table

Semantics

$\lambda N' \lambda Pol \lambda e \lambda x \exists e_a \exists e_n.[Pol(A'(e_a)) \wedge x = e_a \wedge N'(\lambda S.S)(e_n)(x)]$

(e.g. *oval*)

$\lambda N' \lambda Pol \lambda e \lambda x \exists e_a \exists e_n.[Pol(\textsf{oval}(e_a)) \wedge x = e_a \wedge N'(\lambda S.S)(e_n)(x)]$

A.1.8 Class Ipa8

Members

artisanal1, bacterial1, biennial1, biologic1, bisyllabic1, constitutional1, cultural1, economic1, immediate4, metallic1, monthly1, philatelic1, weekly, wooden1, yearly1

Features

Semantic Properties			
Subsectivity	Absoluteness	Transparency	
S	A	T	
Lexical Semantic Properties			
Gradability	Very-Mod	Antonymy	Stability
NG	-	M	S
Syntactic Properties			
Pred/Attr Construction	For/As Construction	Complementation	NP_{meas}
PA	-	-	-
Morphoderivational Properties			
N_{rel}			

Inference Patterns

P1: This is an artisanal product $\leftrightarrow$ This product is artisanal
P10: This product is artisanal $\rightarrow$ This product is not machine-made
P13: This is an artisanal product $\rightarrow$ This is an artisanal object
P14: This is an artisanal object and is a table $\rightarrow$ This is an artisanal table
P28: This is an artisanal product $\leftrightarrow$ This product is made by an artisan
P31: This is an artisanal product $\rightarrow$ This is artisanal
P34: This is an artisanal product $\rightarrow$ This is a product

Semantics

$\lambda N' \lambda Pol \lambda e \lambda x \exists e_a \exists e_n.[Pol(A'(e_a)) \wedge Pol(Rel(e_a, x)) \wedge N'(\lambda S.S)(e_n)(x)]$

(e.g. *weekly*)

$\lambda N' \lambda Pol \lambda e \lambda x \exists e_a \exists e_n \exists d.[Pol(\textit{frequency}(e_a)) \wedge \textit{week}(d) \wedge \textit{value}(e_a, d) \wedge \textit{has-p}(e_a, x)$
$\wedge N'(\lambda S.S)(e_n)(x)]$

A.1.9 Class Ipa9

Members

blind1, blond1, brunet1, deaf1, edible1, even1, finite1, fractional1, infinite1, mute3, odd1, partial1, permanent1, spineless1, temporary1, unique3, vacuous3, various1

Features

Semantic Properties			
Subsectivity	Absoluteness	Transparency	
S	A	T	
Lexical Semantic Properties			
Gradability	Very-Mod	Antonymy	Stability
NG	-	B	S
Syntactic Properties			
Pred/Attr Construction	For/As Construction	Complementation	NP_{meas}
PA	-	-	-
Morphoderivational Properties			
N_a			

Inference Patterns

P1: This is a finite set ↔ This set is finite
P10: This set is finite → This set is not infinite
P11: This set is not finite → This set is infinite
P13: This is a finite set → This is a finite entity
P14: This is a finite list and is a list of numbers → This is a finite list of numbers
P25: It is clear that this is a finite set ↔ The finiteness of this set is clear
P31: This is a finite set → This is finite
P34: This is a finite set → This is a set

Semantics

$\lambda N' \lambda Pol \lambda e \lambda x \exists e_a \exists e_n.[Pol(A'(e_a)) \wedge x = e_a \wedge N'(\lambda S.S)(e_n)(x)]$

(e.g. *finite*)

$\lambda N \lambda Pol \lambda e \lambda x \exists e_a \exists e_n.[Pol(\textsf{finite}(e_a)) \wedge x = e_a \wedge N(\lambda S.S)(e_n)(x)]$

A.1.10 Class Ipa10

Members

actual2, authentic2, genuine1, natural2, real1

Features

Semantic Properties					
Subsectivity		Absoluteness		Transparency	
S		A		T	
Lexical Semantic Properties					
Gradability	Very-Mod		Antonymy		Stability
NG	-		B		S
Syntactic Properties					
Pred/Attr Construction	For/As Construction		Complementation		NP_{meas}
PA	-		SubjE(S-only)		-
Morphoderivational Properties					
N_a					

Inference Patterns

P1: This is a real table $\leftrightarrow$ This table is real
P4: It is real that birth defects occur $\leftrightarrow$ That birth defects occur is real
P10: This object is real $\rightarrow$ This object is not unreal
P11: This object is not real $\rightarrow$ This object is unreal
P13: This is a real table $\rightarrow$ This is a real object
P14: This is a real object and is a table $\rightarrow$ This is a real table
P25: It is clear that this table is real $\leftrightarrow$ The realness of the table is clear
P31: This is a real table $\rightarrow$ This is real
P34: This is a real table $\rightarrow$ This is a table

Semantics

$\lambda N' \lambda Pol \lambda e \lambda x \exists e_a \exists e_n.[Pol(A'(e_a)) \wedge x = e_a \wedge N'(\lambda S.S)(e_n)(x)]$

(e.g. *genuine*)

$\lambda N \lambda Pol \lambda e \lambda x \exists e_a \exists e_n.[Pol(\textsf{genuine}(e_a)) \wedge x = e_a \wedge N(\lambda S.S)(e_n)(x)]$

A.1.11 Class Ipa11

Members

accompanied1, barred1, bent1, broken1, closed1, clothed1, cooked1, crooked1, damaged1, dead1, dressed1, fastened1, hand-made1, locked1, machine-made1, married1, open1, scared1, shut1, unbroken1, unclothed1, undamaged1, unfastened1, unmarried1

Features

Semantic Properties			
Subsectivity	Absoluteness	Transparency	
S	A	T	
Lexical Semantic Properties			
Gradability	Very-Mod	Antonymy	Stability
NG	-	B	D
Syntactic Properties			
Pred/Attr Construction	For/As Construction	Complementation	NP$_{meas}$
PA	-	-	-
Morphoderivational Properties			
V$_{res}$			

Inference Patterns

P1: This is a closed door ↔ This door is closed
P10: This door is closed → This door is not open
P11: This door is not closed → This door is open
P13: This is a closed door → This is a closed object
P14: This is a closed object and is a door → This is a closed door
P20: Someone has closed the door → The door is closed
P31: This is a closed door → This is closed
P34: This is a closed door → This is a door

Semantics

$\lambda Pol \lambda N' \lambda e \lambda x \exists e_a \exists e_n.[\textsf{\textit{V}}_a(e_a) \wedge Pol(\textsf{\textit{patient}}(e_a, x)) \wedge N'(\lambda S.S)(e_n)(x)]$

(e.g. *closed*)

$\lambda Pol \lambda N' \lambda e \lambda x \exists e_a \exists e_n.[\textsf{\textit{close}}(e_a) \wedge Pol(\textsf{\textit{patient}}(e_a, x)) \wedge N'(\lambda S.S)(e_n)(x)]$

A.1.12 Class Ipa12

Members

absent1, dynamic3, motionless1, omnipresent1, present2, static1, ubiquitous1

Features

Semantic Properties			
Subsectivity	Absoluteness	Transparency	
S	A	T	
Lexical Semantic Properties			
Gradability	Very-Mod	Antonymy	Stability
NG	-	B	D
Syntactic Properties			
Pred/Attr Construction	For/As Construction	Complementation	NP$_{meas}$
PA	-	-	-
Morphoderivational Properties			
N$_a$			

Inference Patterns

P10: John is present $\rightarrow$ John is not absent
P11: John is not present $\rightarrow$ John is absent
P13: This is a present student $\rightarrow$ This is a present person
P14: This is a present person and is a student $\rightarrow$ This is a present student
P25: It is clear that John is present $\leftrightarrow$ The presence of John is clear
P31: John is a present student $\rightarrow$ John is present
P34: John is a present student $\rightarrow$ John is a student

Semantics

$\lambda N' \lambda Pol \lambda e \lambda x \exists e_a \exists e_n.[Pol(A'(e_a)) \wedge x = e_a \wedge N'(\lambda S.S)(e_n)(x)]$

(e.g. *absent*)

$\lambda N \lambda Pol \lambda e \lambda x \exists e_a \exists e_n.[Pol(\textit{absent}(e_a)) \wedge x = e_a \wedge N(\lambda S.S)(e_n)(x)]$

A.1.13 Class Ip1

Members

ablaze2, afire1, aflame2, aglow1, afraid1, alive1, asleep1, awake1

Features

Semantic Properties			
Subsectivity	Absoluteness	Transparency	
S	A	T	
Lexical Semantic Properties			
Gradability	Very-Mod	Antonymy	Stability
G	-	B	D
Syntactic Properties			
Pred/Attr Construction	For/As Construction	Complementation	NP_{meas}
P	-	-	-
Morphoderivational Properties			
V_a, N_e			

Inference Patterns

P10: John is asleep → John is not awake
P11: John is not asleep → John is awake
P13: This man is asleep → This entity is asleep
P14: This is an asleep man and is a student → This student is asleep
P17: John is asleep ↔ John sleeps
P23: John is asleep ↔ John is the sleeper
P24: John is deeply asleep ↔ John's sleep is deep
P31: This man is asleep → This is asleep
P34: This man is asleep → This is a man

Semantics

$\lambda Pol \lambda N' \lambda e \lambda x \exists e_a \exists e_n \exists a \exists d.[\mathsf{V}_a(e_a) \wedge \textit{extend}(e_a, d) \wedge \textit{absStandard}(a, v_a) \wedge Pol(\textit{theme}(e_a, x)) \wedge N'(\lambda S.S)(e_n)(x)]$

(e.g. *asleep*)

$\lambda Pol \lambda N' \lambda e \lambda x \exists e_a \exists e_n \exists a \exists d.[\textit{sleep}(e_a) \wedge \textit{extend}(e_a, d) \wedge \textit{absStandard}(a, sleep) \wedge Pol(\textit{theme}(e_a, x)) \wedge N'(\lambda S.S)(e_n)(x)]$

A.1.14 Class Ip2

Members

adrift2, afloat2, aground1

Features

Semantic Properties			
Subsectivity	Absoluteness	Transparency	
S	A	T	
Lexical Semantic Properties			
Gradability	Very-Mod	Antonymy	Stability
NG	-	M	D
Syntactic Properties			
Pred/Attr Construction	For/As Construction	Complementation	NP$_{meas}$
P	-	-	-
Morphoderivational Properties			
V$_a$, N$_e$			

Inference Patterns

P10: The boat is afloat $\rightarrow$ This boat is not aground
P13: The boat is afloat $\rightarrow$ This object is afloat
P14: This object is afloat and is a boat $\rightarrow$ This boat is afloat
P17: This boat is afloat $\leftrightarrow$ This boat floats
P23: This boat is afloat $\leftrightarrow$ This boot is the floater
P24: The boat is slowly afloat $\leftrightarrow$ The boat floating is slowly
P31: The boat is afloat $\rightarrow$ This is afloat
P34: The boat is afloat $\rightarrow$ This is a boat

Semantics

$\lambda Pol \lambda N' \lambda e \lambda x \exists e_a \exists e_n.[\mathit{V}_a(e_a) \wedge Pol(\mathit{theme}(e_a, x)) \wedge N'(\lambda S.S)(e_n)(x)]$

(e.g. *afloat*)

$\lambda Pol \lambda N' \lambda e \lambda x \exists e_a \exists e_n.[\mathit{float}(e_a) \wedge Pol(\mathit{theme}(e_a, x)) \wedge N'(\lambda S.S)(e_n)(x)]$

A.1.15 Class Ia1

Members

present1

Features

Semantic Properties			
Subsectivity	Absoluteness	Transparency	
S	A	T	
Lexical Semantic Properties			
Gradability	Very-Mod	Antonymy	Stability
NG	-	M	S
Syntactic Properties			
Pred/Attr Construction	For/As Construction	Complementation	NP$_{meas}$
A	-	-	-
Morphoderivational Properties			
N$_{rel}$			

Inference Patterns

P10: John is the present president $\rightarrow$ John is not the past president
P13: John is the present president $\rightarrow$ John is a present entity
P14: John is a present entity and is a president $\rightarrow$ John is a present president
P28: John is the present president $\leftrightarrow$ John is the president at the present time
P34: John is the present president $\rightarrow$ John is a president

Semantics

$\lambda N' \lambda Pol \lambda e \lambda x \exists e_a \exists e_n.[Pol(A'(e_a)) \wedge \textit{time}(e_a, x) \wedge N'(\lambda S.S)(e_n)(x)]$

(e.g. *present*)

$\lambda N \lambda Pol \lambda e \lambda x \exists e_a \exists e_n.[Pol(\textit{present}(e_a)) \wedge \textit{time}(e_a, x) \wedge N(\lambda S.S)(e_n)(x)]$

A.2 Subsective Adjectives

A.2.1 Class Spa1

Members

atomic3, average1, enormous1, giant1, immense1, intermediate2, medium1, oceanic2, superb2, tepid1

Features

Semantic Properties			
Subsectivity	Absoluteness	Transparency	
S	NA	T	
Lexical Semantic Properties			
Gradability	Very-Mod	Antonymy	Stability
NG	-	C	S
Syntactic Properties			
Pred/Attr Construction	For/As Construction	Complementation	NP_{meas}
PA	for	-	-
Morphoderivational Properties			
N_a			

Inference Patterns

P1: This is an enormous mouse $\leftrightarrow$ This mouse is enormous
P2: X is an enormous mouse $\leftrightarrow$ X is enormous for a mouse
P10: This mouse is enormous $\rightarrow$ This mouse is not small
P25: It is clear that this mouse is enormous $\leftrightarrow$ The enormousness of the mouse is clear
P29: The music was enormously amplified $\leftrightarrow$ The amplification of the music was enormous
P34: This mouse is enormous $\rightarrow$ This is a mouse

Semantics

$\lambda Pol \lambda N' \lambda e \lambda x.\exists e_a \exists e_n \exists a \exists d[\textsf{N}_a(e_a) \wedge \textsf{has-p}(e_a, x) \wedge \textsf{extent}(e_a, d) \wedge \textsf{standard}(a, n) \wedge Pol(\textsf{max}/\textsf{min}/\textsf{average}(d, a)) \wedge N'(\lambda S.S)(e_n)(x))]$

(e.g. *enormous*)

$\lambda N' \lambda Pol \lambda e \lambda x.\exists e_a \exists e_n \exists a \exists d[\textsf{size}(e_a) \wedge Pol(\textsf{has-p}(e_a, x)) \wedge \textsf{extent}(e_a, d) \wedge \textsf{standard}(a, n) \wedge Pol(\textsf{max}(d, a)) \wedge N'(\lambda S.S)(e_n)(x))]$

A.2.2 Class Spa2

Members

aged1, big1, deep3, fat1, heavy1, high1, huge1, large1, long1, narrow1, old1, tall1, thick1, wide1

Features

Semantic Properties			
Subsectivity	Absoluteness	Transparency	
S	NA	T	
Lexical Semantic Properties			
Gradability	Very-Mod	Antonymy	Stability
G	+	C	S
Syntactic Properties			
Pred/Attr Construction	For/As Construction	Complementation	NP_{meas}
PA	for	-	+
Morphoderivational Properties			
N_a			

Inference Patterns

P1: This is a big mouse ↔ This mouse is big
P2: X is a big mouse ↔ X is big for a mouse
P10: This mouse is big → This mouse is not small
P25: It is clear that this mouse is big ↔ The bigness of the mouse is clear
P33: John is a 2 meter tall footballer → John is 2 meter tall
P34: This mouse is big → This is a mouse

Semantics

$\lambda Pol \lambda N' \lambda e \lambda x.\exists e_a \exists e_n \exists a \exists d[\textsf{N}_a(e_a) \wedge \textit{has-p}(e_a, x) \wedge \textit{extent}(e_a, d) \wedge Pol(\textit{highFor}(d, a)) \wedge \textit{standard}(a, n) \wedge N'(\lambda S.S)(e_n)(x))]$

(e.g. *tall*)

$\lambda N' \lambda Pol \lambda e \lambda x.\exists e_a \exists e_n \exists a \exists d[\textit{size}(e_a) \wedge Pol(\textit{has-p}(e_a, x)) \wedge \textit{extent}(e_a, d) \wedge Pol(\textit{highFor}(d, a)) \wedge \textit{standard}(a, n) \wedge \wedge N'(\lambda S.S)(e_n)(x))]$

A.2.3 Class Spa3

Members

colossal1, gigantic1, great1, hard3, light1, little1, low1, miniature1, minuscule3, monumental1, shallow1, short3, short1, small1, soft1, thin2, vast1, young1

Features

Semantic Properties			
Subsectivity	Absoluteness	Transparency	
S	NA	T	
Lexical Semantic Properties			
Gradability	Very-Mod	Antonymy	Stability
G	+	C	S
Syntactic Properties			
Pred/Attr Construction	For/As Construction	Complementation	NP$_{meas}$
PA	for	-	-
Morphoderivational Properties			
N$_a$			

Inference Patterns

P1: This is a small mouse ↔ This mouse is small
P2: X is a small mouse ↔ X is small for a mouse
P10: This mouse is small → This mouse is not big
P25: It is clear that this mouse is small ↔ The smallness of the mouse is clear
P34: This mouse is small → This is a mouse

Semantics

$\lambda Pol \lambda N' \lambda e \lambda x.\exists e_a \exists e_n \exists a \exists d[\textsf{N}_a(e_a) \wedge \textsf{has-p}(e_a, x) \wedge \textsf{extent}(e_a, d) \wedge Pol(\textsf{lowFor}(d, a)) \wedge \textsf{standard}(a, n) \wedge \wedge N'(\lambda S.S)(e_n)(x))]$

(e.g. *small*)

$\lambda N' \lambda Pol \lambda e \lambda x.\exists e_a \exists e_n \exists a \exists d[\textsf{size}(e_a) \wedge Pol(\textsf{has-p}(e_a, x)) \wedge \textsf{extent}(e_a, d) \wedge Pol(\textsf{lowFor}(d, a)) \wedge \textsf{standard}(a, n) \wedge N'(\lambda S.S)(e_n)(x))]$

A.2.4 Class Spa4

Members

expensive1, fast1, hot1, late2, loud1, quick1, rich1, strong1, warm1, wealthy1

Features

Semantic Properties			
Subsectivity	Absoluteness	Transparency	
S	NA	T	
Lexical Semantic Properties			
Gradability	Very-Mod	Antonymy	Stability
G	+	C	D
Syntactic Properties			
Pred/Attr Construction	For/As Construction	Complementation	NP$_{meas}$
PA	for	-	+
Morphoderivational Properties			
N$_a$			

Inference Patterns

P1: This is a quick horse $\leftrightarrow$ This horse is quick
P2: X is a quick horse $\leftrightarrow$ X is quick for a horse
P10: This horse is quick $\rightarrow$ This horse is not slow
P22: This is a fast car $\leftrightarrow$ This car runs fast
P25: It is clear that this horse is quick $\leftrightarrow$ The quickness of the horse is clear
P29: John's answer was quick $\leftrightarrow$ John answered quickly
P33: This is a 2 dollar expensive book $\rightarrow$ This is 2 dollar expensive
P34: This horse is quick $\rightarrow$ This is a horse

Semantics

$\lambda Pol \lambda N' \lambda e \lambda x.\exists e_a \exists e_n \exists a \exists d[\textsf{N}_a(e_a) \wedge \textit{has-p}(e_a, x) \wedge \textit{extent}(e_a, d) \wedge Pol(\textit{highFor}(d, a)) \wedge \textit{standard}(a, n) \wedge \wedge N'(\lambda S.S)(e_n)(x))]$

(e.g. *fast*)

$\lambda N' \lambda Pol \lambda e \lambda x.\exists e_a \exists e_n \exists a \exists d[\textit{speed}(e_a) \wedge Pol(\textit{has-p}(e_a, x)) \wedge \textit{extent}(e_a, d) \wedge Pol(\textit{highFor}(d, a)) \wedge \textit{standard}(a, n) \wedge \wedge N'(\lambda S.S)(e_n)(x))]$

A.2.5 Class Spa5

Members

cheap1, cold1, cool1, icy2, gelid1, poor2, shrill1, slow1, soft3, swift1, weak1

Features

Semantic Properties			
Subsectivity	Absoluteness	Transparency	
S	NA	T	
Lexical Semantic Properties			
Gradability	Very-Mod	Antonymy	Stability
G	+	C	D
Syntactic Properties			
Pred/Attr Construction	For/As Construction	Complementation	NP$_{meas}$
PA	for	-	-
Morphoderivational Properties			
N$_a$			

Inference Patterns

P1: This is a slow horse $\leftrightarrow$ This horse is slow
P2: X is a slow horse $\leftrightarrow$ X is slow for a horse
P10: This horse is slow $\rightarrow$ This horse is not fast
P25: It is clear that this horse is slow $\leftrightarrow$ The slowness of the horse is clear
P29: John's answer was slow $\leftrightarrow$ John answered slowly
P34: This horse is slow $\rightarrow$ This is a horse

Semantics

$\lambda Pol \lambda N' \lambda e \lambda x.\exists e_a \exists e_n \exists a \exists d[\textit{N}_a(e_a) \wedge \textit{has-p}(e_a, x) \wedge \textit{extent}(e_a, d) \wedge Pol(\textit{lowFor}(d, a)) \wedge \textit{standard}(a, n) \wedge \wedge N'(\lambda S.S)(e_n)(x))]$

(e.g. *slow*)

$\lambda N' \lambda Pol \lambda e \lambda x.\exists e_a \exists e_n \exists a \exists d[\textit{speed}(e_a) \wedge Pol(\textit{has-p}(e_a, x)) \wedge \textit{extent}(e_a, d) \wedge Pol(\textit{lowFor}(d, a)) \wedge \textit{standard}(a, n) \wedge N'(\lambda S.S)(e_n)(x))]$

A.2.6 Class Spa6

Members

dangerous1, pleasant1, profitable1, safe1, unpleasant1, unprofitable1, useful1, useless1

Features

Semantic Properties			
Subsectivity	Absoluteness	Transparency	
S	NA	NT	
Lexical Semantic Properties			
Gradability	Very-Mod	Antonymy	Stability
G	+	B	D
Syntactic Properties			
Pred/Attr Construction	For/As Construction	Complementation	NP$_{meas}$
PA	as	SubjE(Easy)	-
Morphoderivational Properties			
N$_a$			

Inference Patterns

P1: This is a useful task $\leftrightarrow$ This task is useful
P3: This is a useful task $\leftrightarrow$ This is useful as a task
P4: It is useful to work hard $\leftrightarrow$ To work hard is useful
P8: John is useful to talk to $\leftrightarrow$ It is useful to talk to John
P9: John is useful for Mary to talk to $\leftrightarrow$ It is useful for Mary to talk to John
P10: This task is useful $\rightarrow$ This task is not useless
P11: This task is not useful $\rightarrow$ This task is useless
P25: It is clear that this knife is useful $\leftrightarrow$ The usefulness of this knife is clear
P34: This is a useful knife $\rightarrow$ This is a knife

Semantics

$\lambda N' \lambda Pol \lambda e \lambda x. \exists e_a \exists e_n [\textsf{N}_a(e_a) \wedge Pol(\textsf{sem-role}(e_a, x)) \wedge \textsf{role\_as}(e_a, n) \wedge \textsf{N'}(\lambda S.S)(e_n)(x)]$

(e.g. *dangerous*)

$\lambda N' \lambda Pol \lambda e \lambda x. \exists e_a \exists e_n [\textsf{danger}(e_a) \wedge Pol(\textsf{sem-role}(e_a, x)) \wedge \textsf{role\_as}(e_a, n) \wedge \textsf{N'}(\lambda S.S)(e_n)(x)]$

A.2.7 Class Spa7

Members

certain2, frequent1, habitual1, infrequent1, rare2, sure1, unusual1, usual1

Features

Semantic Properties					
Subsectivity		Absoluteness		Transparency	
S		NA		NT	
Lexical Semantic Properties					
Gradability	Very-Mod		Antonymy		Stability
G	+		B		D
Syntactic Properties					
Pred/Attr Construction	For/As Construction		Complementation		NP$_{meas}$
PA	as		SubjE(for)		-
Morphoderivational Properties					
N$_a$					

Inference Patterns

P1: This is a certain answer ↔ This answer is certain
P3: This is a certain answer ↔ This is certain as an answer
P4: It is certain that John will visit us ↔ That John will visit us is certain
P6: John is certain to visit us ↔ It is certain for John to visit us
P10: This answer is certain → This answer is not uncertain
P11: This answer is not certain → This answer is uncertain
P25: It is clear that this answer is certain ↔ The certainty of this answer is clear
P34: This answer is certain → This is an answer

Semantics

$\lambda N' \lambda Pol \lambda e \lambda x.\exists e_a \exists e_n[\textsf{N}_a(e_a) \wedge Pol(\textsf{\textit{sem-role}}(e_a, x)) \wedge \textsf{\textit{role\_as}}(e_a, n) \wedge \textsf{\textit{N'}}(\lambda S.S)(e_n)(x)]$

(e.g. *certain*)

$\lambda N' \lambda Pol \lambda e \lambda x.\exists e_a \exists e_n[\textsf{\textit{certainty}}(e_a) \wedge Pol(\textsf{\textit{sem-role}}(e_a, x)) \wedge \textsf{\textit{role\_as}}(e_a, n) \wedge \textsf{\textit{N'}}(\lambda S.S)(e_n)(x)]$

A.2.8 Class Spa8

Members

acknowledged1, appropriate1, arbitrary1, damaging1, decisive1, famous1, known1, inappropriate1, indecisive2, insufficient1, necessary1, scarce1, sufficient1, true1, unacknowledged1, uncontroversial1, unknown1, unnecessary1

Features

Semantic Properties			
Subsectivity	Absoluteness	Transparency	
S	NA	NT	
Lexical Semantic Properties			
Gradability	Very-Mod	Antonymy	Stability
G	+	B	D
Syntactic Properties			
Pred/Attr Construction	For/As Construction	Complementation	NP_{meas}
PA	as	SubjE(S-only)	-
Morphoderivational Properties			
N_a			

Inference Patterns

P1: This is a necessary task ↔ This task is necessary
P3: This is a necessary task ↔ This is necessary as a task
P4: It is necessary to work hard ↔ To work hard is necessary
P10: This answer is necessary → This answer is not unnecessary
P11: This answer is not necessary → This answer is unnecessary
P25: It is clear that John's work is necessary ↔ The necessity of John's work is clear
P34: This is a necessary task → This is a task

Semantics

$\lambda N' \lambda Pol \lambda e \lambda x.\exists e_a \exists e_n[\textsf{N}_a(e_a) \wedge Pol(\textsf{sem-role}(e_a, x)) \wedge \textsf{role\_as}(e_a, n) \wedge \textsf{N'}(\lambda S.S)(e_n)(x)]$

(e.g. *necessary*)

$\lambda N' \lambda Pol \lambda e \lambda x.\exists e_a \exists e_n[\textsf{necessity}(e_a) \wedge Pol(\textsf{necessee}(e_a, x)) \wedge \textsf{role\_as}(e_a, n) \wedge \textsf{N'}(\lambda S.S)(e_n)(x)$

A.2.9 Class Spa9

Members

clear1, evident1, new1, novel2, obscure1, old2, obvious1, recent1, unclear1

Features

Semantic Properties			
Subsectivity	Absoluteness	Transparency	
S	NA	NT	
Lexical Semantic Properties			
Gradability	Very-Mod	Antonymy	Stability
G	+	B	D
Syntactic Properties			
Pred/Attr Construction	For/As Construction	Complementation	NP$_{meas}$
PA	as	SubjE(to)	-
Morphoderivational Properties			
N$_a$			

Inference Patterns

P1: This is a clear answer ↔ This answer is clear
P3: This is a clear answer ↔ This is clear as an answer
P4: It is clear that it will rain ↔ That it will rain is clear
P7: Sam is clear that it will rain ↔ It is clear to Sam that it will rain
P10: This answer is clear → This answer is not unclear
P11: This answer is not clear → This answer is unclear
P25: It is evident that this is a clear answer ↔ The clarity of the answer is evident
P29: John's talk was clear ↔ John talked clearly
P34: This is a clear answer → This is an answer

Semantics

$\lambda N' \lambda Pol \lambda e \lambda x.\exists e_a \exists e_n[\textsf{N}_a(e_a) \wedge Pol(\textsf{sem-role}(e_a, x)) \wedge \textsf{role\_as}(e_a, n) \wedge \textsf{N'}(\lambda S.S)(e_n)(x)]$

(e.g. *clear*)

$\lambda N' \lambda Pol \lambda e \lambda x.\exists e_a \exists e_n[\textit{clarity}(e_a) \wedge Pol(\textit{theme}(e_a, x)) \wedge \textsf{role\_as}(e_a, n) \wedge \textsf{N'}(\lambda S.S)(e_n)(x)]$

A.2.10 Class Spa10

Members

angry1, anxious2, desperate1, disinclined1, eager1, experienced1, impatient1, inexperienced1, jealous2, patient1, ready1, unangry1, uneager1, unready1, unwilling1, willing1

Features

Semantic Properties			
Subsectivity	Absoluteness	Transparency	
S	NA	NT	
Lexical Semantic Properties			
Gradability	Very-Mod	Antonymy	Stability
G	+	B	D
Syntactic Properties			
Pred/Attr Construction	For/As Construction	Complementation	NP$_{meas}$
PA	as	ObjE	-
Morphoderivational Properties			
N$_a$			

Inference Patterns

P1: This is an anxious mother ↔ This mother is anxious
P3: She is an anxious mother ↔ She is anxious as a mother
P10: This is an anxious mother → This mother is not untroubled
P11: This mother is not anxious → This mother is untroubled
P25: It is clear that this is an anxious mother ↔ The anxiousness of this mother is clear
P29: John's behaviour was anxious ↔ John behaved anxiously
P34: This is an anxious person → This is a person

Semantics

$\lambda N' \lambda Pol \lambda e \lambda x . \exists e_a \exists e_n [\mathsf{N}_a(e_a) \wedge Pol(\textsf{sem-role}(e_a, x)) \wedge \textsf{role\_as}(e_a, n) \wedge \textsf{N'}(\lambda S.S)(e_n)(x)]$

(e.g. *angry*)

$\lambda N' \lambda Pol \lambda e \lambda x . \exists e_a \exists e_n [\textsf{anger}(e_a) \wedge Pol(\textsf{experiencer}(e_a, x)) \wedge \textsf{role\_as}(e_a, n) \wedge \textsf{N'}(\lambda S.S)(e_n)(x)]$

A.2.11 Class Spa11

Members

atypical1, awkward2, beautiful1, clumsy1, graceful1, handsome1, hopeful1, hopeless1, intelligent1, pretty1, serious1, skilled1, skillful1, unskilled1, typical1, ugly1, unexpected1

Features

Semantic Properties			
Subsectivity	Absoluteness	Transparency	
S	NA	NT	
Lexical Semantic Properties			
Gradability	Very-Mod	Antonymy	Stability
G	+	C	S
Syntactic Properties			
Pred/Attr Construction	For/As Construction	Complementation	NP_{meas}
PA	as	SubjE(S-only)	-
Morphoderivational Properties			
N_a			

Inference Patterns

P1: This is an intelligent mathematician $\leftrightarrow$ This mathematician is intelligent
P3: Mary is an intelligent mathematician $\leftrightarrow$ Mary is intelligent as a mathematician
P4: It is intelligent to say that $\leftrightarrow$ To say that is intelligent
P10: Sam is intelligent $\rightarrow$ Sam is not stupid
P25: It is evident that this is an intelligent mathematician $\leftrightarrow$ The intelligence of the mathematician is evident
P29: John's talk was intelligent $\leftrightarrow$ John talked intelligently
P34: This is an intelligent mathematician $\rightarrow$ This is an mathematician

Semantics

$\lambda N'\lambda Pol\lambda e\lambda x.\exists e_a\exists e_n[\textit{N}_a(e_a) \wedge Pol(\textit{sem-role}(e_a, x)) \wedge \textit{role\_as}(e_a, n) \wedge \textit{N'}(\lambda S.S)(e_n)(x)]$

(e.g. *intelligent*)

$\lambda N'\lambda Pol\lambda e\lambda x.\exists e_a\exists e_n[\textit{intelligence}(e_a) \wedge Pol(\textit{theme}(e_a, x)) \wedge \textit{role\_as}(e_a, n) \wedge \textit{N'}(\lambda S.S)(e_n)(x)]$

A.2.12 Class Spa12

Members

amusing2, bad1, comfortable1, complex1, difficult1, easy1, excellent1, exciting1, funny1, good1, hard1, important1, interesting1, nice1, perfect1, satisfying1, simple1, tough2, tricky2, unamusing1, uncomfortable1, unexciting2, unimportant1, uninteresting1, unsatisfying1

Features

Semantic Properties			
Subsectivity	Absoluteness	Transparency	
S	NA	NT	
Lexical Semantic Properties			
Gradability	Very-Mod	Antonymy	Stability
G	+	C	D
Syntactic Properties			
Pred/Attr Construction	For/As Construction	Complementation	NP_{meas}
PA	as	SubjE(Easy)	-
Morphoderivational Properties			
N_a			

Inference Patterns

P1: This is an easy exercise $\leftrightarrow$ This exercise is easy
P3: This is a easy task $\leftrightarrow$ This is easy as a task
P4: It is easy to solve the problem $\leftrightarrow$ To solve the problem is easy
P8: John is easy to talk to $\leftrightarrow$ It is easy to talk to John
P9: John is easy for Mary to talk to $\leftrightarrow$ It is easy for Mary to talk to John
P10: This task is easy $\rightarrow$ This task is not difficult
P25: It is clear that this problem is easy $\leftrightarrow$ The easiness of this problem is clear
P34: This is an easy answer $\rightarrow$ This is an answer

Semantics

$\lambda N' \lambda Pol \lambda e \lambda x. \exists e_a \exists e_n [\mathsf{N}_a(e_a) \wedge Pol(\textsf{sem-role}(e_a, x)) \wedge \textsf{role\_as}(e_a, n) \wedge \textsf{N'}(\lambda S.S)(e_n)(x)]$

(e.g. *easy*)

$\lambda N' \lambda Pol \lambda e \lambda x. \exists e_a \exists e_n [\textsf{difficulty}(e_a) \wedge Pol(\textsf{activity}(e_a, x)) \wedge \textsf{role\_as}(e_a, n) \wedge \textsf{N'}(\lambda S.S)(e_n)(x)]$

A.2.13 Class Spa13

Members

altruistic1, ambitious1, artful1, brave1, capricious1, civil2, clever3, considerate1, crazy1, criminal3, cruel1, disloyal2, egoistic1, foolish1, friendly1, generous1, humane1, immoral1, impolite1, inconsiderate1, ingenuous1, inhumane1, insane1, insensible1, just1, kind1, loyal3, polite1, rude1, sensible1, stingy1, stupid1, tender1, tough1, unfriendly1, unjust2, unkind1, unreasonable1, wise1

Features

Semantic Properties			
Subsectivity	Absoluteness	Transparency	
S	NA	NT	
Lexical Semantic Properties			
Gradability	Very-Mod	Antonymy	Stability
G	+	C	D
Syntactic Properties			
Pred/Attr Construction	For/As Construction	Complementation	NP$_{meas}$
PA	as	SubjE(of)	-
Morphoderivational Properties			
N$_a$			

Inference Patterns

P1: He is a cruel father ↔ This father is cruel
P3: He is a cruel father ↔ He is cruel as a father
P4: It is cruel to say that ↔ To say that is cruel
P5: John is cruel to say that ↔ To say that is cruel of John
P10: This answer is cruel → This answer is not humane
P14: John is a cruel man → John is a cruel husband
P15: It is cruel to say that → It is cruel to say that loud
P25: It is clear that John is cruel ↔ The cruelty of John is clear
P29: John's smile was cruel ↔ John smiled cruelly
P34: This is a cruel husband → This is a husband

Semantics

$\lambda N' \lambda Pol \lambda e \lambda x.\exists e_a \exists e_n[\textsf{N}_a(e_a) \wedge Pol(\textsf{sem-role}(e_a, x)) \wedge \textsf{role\_as}(e_a, n) \wedge \textsf{N'}(\lambda S.S)(e_n)(x)]$

(e.g. *cruel*)

$\lambda N' \lambda Pol \lambda e \lambda x.\exists e_a \exists e_n[\textsf{cruelty}(e_a) \wedge Pol(\textsf{theme}(e_a, x)) \wedge \textsf{role\_as}(e_a, n) \wedge \textsf{N'}(\lambda S.S)(e_n)(x)]$

A.2.14 Class Spa14

Members

fortunate1, happy1, sad1, unhappy1, unfortunate1

Features

Semantic Properties			
Subsectivity	Absoluteness	Transparency	
S	NA	NT	
Lexical Semantic Properties			
Gradability	Very-Mod	Antonymy	Stability
G	+	C	D
Syntactic Properties			
Pred/Attr Construction	For/As Construction	Complementation	NP_{meas}
PA	as	SubjE(for)	-
Morphoderivational Properties			
N_a			

Inference Patterns

P1: This is a sad answer $\leftrightarrow$ This answer is sad
P3: This is a sad answer $\leftrightarrow$ This is sad as an answer
P4: It is sad that you go $\leftrightarrow$ That you go is sad
P6: I'm sad to leave $\leftrightarrow$ To leave is sad for me
P10: This man is sad $\rightarrow$ This man is not glad
P14: John is a sad man $\rightarrow$ John is a sad student
P15: It is sad to say that $\rightarrow$ It is sad to say that loud
P25: It is clear that John is sad $\leftrightarrow$ The sadness of John is clear
P29: John's smile was sad $\leftrightarrow$ John smiled sadly
P34: John is a sad student $\rightarrow$ John is a student

Semantics

$\lambda N' \lambda Pol \lambda e \lambda x.\exists e_a \exists e_n[\textit{N}_a(e_a) \wedge Pol(\textit{sem-role}(e_a, x)) \wedge \textit{role\_as}(e_a, n) \wedge \textit{N'}(\lambda S.S)(e_n)(x)]$

(e.g. *sad*)

$\lambda N' \lambda Pol \lambda e \lambda x.\exists e_a \exists e_n[\textit{sadness}(e_a) \wedge Pol(\textit{experiencer}(e_a, x)) \wedge \textit{role\_as}(e_a, n) \wedge \textit{N'}(\lambda S.S)(e_n)(x)]$

A.2.15 Class Sp1

Members

averse1, aware1, fond4, inclined1, indisposed2, disinclined1, loath2, reluctant1

Features

Semantic Properties			
Subsectivity	Absoluteness	Transparency	
S	NA	NT	
Lexical Semantic Properties			
Gradability	Very-Mod	Antonymy	Stability
G	+	C	D
Syntactic Properties			
Pred/Attr Construction	For/As Construction	Complementation	NP_{meas}
P	as	PP, ObjE	-
Morphoderivational Properties			
N_a			

Inference Patterns

P3: This student is reluctant to admit mistakes $\leftrightarrow$ As a student he is reluctant to admit mistakes

P10: This student is reluctant to admit mistakes $\rightarrow$ This student is not willing to admit mistakes

P11: This student is not reluctant to admit mistakes $\rightarrow$ This student is willing to admit mistakes

P25: It is clear that John is reluctant to admit mistakes $\leftrightarrow$ The reluctance of John to admit mistakes is clear

P34: This man is loath to admit mistakes $\rightarrow$ This is a man

Semantics

$\lambda N' \lambda Pol \lambda e \lambda x. \exists e_a \exists e_n [\textsf{N}_a(e_a) \wedge Pol(\textsf{sem-role}(e_a, x)) \wedge \textsf{role\_as}(e_a, n) \wedge \textsf{N'}(\lambda S.S)(e_n)(x)]$

(e.g. *disinclined*)

$\lambda N' \lambda Pol \lambda e \lambda x. \exists e_a \exists e_n [\textit{inclination}(e_a) \wedge Pol(\textit{experiencer}(e_a, x)) \wedge \textsf{role\_as}(e_a, n) \wedge \textsf{N'}(\lambda S.S)(e_n)(x)]$

A.2.16 Class Sa1

Members

acid3, animal1, aquatic1, atomic1, chemical1, civil1, criminal0, dental1, electrical1, feminine1, financial1, gastronomical1, linguistic1, lunar1, marginal1, marine1, masculine1, maternal1, mathematical1, mechanical3, medical1, metal1, molecular1, moral1, musical1, national1, nuclear2, oceanic1, paternal1, plastic1, polar4, presidential1, provincial1, solar1, stellar2, urban1, viral1

Features

Semantic Properties			
Subsectivity	Absoluteness	Transparency	
S	NA	T	
Lexical Semantic Properties			
Gradability	Very-Mod	Antonymy	Stability
NG	-	M	S
Syntactic Properties			
Pred/Attr Construction	For/As Construction	Complementation	NP$_{meas}$
A	-	-	-
Morphoderivational Properties			
N$_{rel}$			

Inference Patterns

P14: This is a solar devise and is a generator $\rightarrow$ This is a solar generator
P21: John is a linguistic teacher $\leftrightarrow$ John teaches linguistics
P28: This is a gastronomical dictionary $\leftrightarrow$ This is a dictionary about gastronomy
P34: This is a gastronomical dictionary $\rightarrow$ This is a dictionary

Semantics

$\lambda N' \lambda Pol \lambda e \lambda x \exists e_a \exists e_n.[N_a(e_a) \wedge Pol(\textsf{\textit{Rel}}(e_a, e_n)) \wedge N'(\lambda S.S)(e_n)(x)]$

(e.g. *gastronomical*)

$\lambda N' \lambda Pol \lambda e \lambda x \exists e_a \exists e_n.[\textsf{\textit{gastronomy}}(e_a) \wedge Pol(\textsf{\textit{about}}(e_a, e_n)) \wedge N'(\lambda S.S)(e_n)(x)]$

A.2.17 Class Sa2

Members

main1, primary3, principal1

Features

Semantic Properties			
Subsectivity	Absoluteness	Transparency	
S	NA	NT	
Lexical Semantic Properties			
Gradability	Very-Mod	Antonymy	Stability
NG	-	C	D
Syntactic Properties			
Pred/Attr Construction	For/As Construction	Complementation	NP$_{meas}$
A	-	-	-
Morphoderivational Properties			

Inference Patterns

P34: This is the main problem → This is a problem

Semantics

$\lambda N' \lambda Pol \lambda e \lambda x.\exists e_a \exists e_n \exists d[\textit{N}_a(e_a) \wedge \textit{ranking}(e_a, d) \wedge \textit{role\_as}(e_a, n) \wedge Pol(\textit{value}(d, adj)) \wedge \textit{N'}(\lambda S.S)(e_n)(x)]$

(e.g. *main*)

$\lambda N' \lambda Pol \lambda e \lambda x.\exists e_a \exists e_n \exists d[\textit{importance}(e_a) \wedge \textit{ranking}(e_a, d) \wedge \textit{role\_as}(e_a, n) \wedge Pol(\textit{value}(d, \textit{main})) \wedge \textit{N'}(\lambda S.S)(e_n)(x)]$

A.2.18 Class Sa3

Members

complete4, consummate3, inveterate1, light17, mere1, perfect2, sound8, utter1, veteran1, whole1

Features

Semantic Properties					
Subsectivity		Absoluteness		Transparency	
S		NA		NT	
Lexical Semantic Properties					
Gradability	Very-Mod		Antonymy		Stability
NG	-		C		S
Syntactic Properties					
Pred/Attr Construction	For/As Construction		Complementation		NP_{meas}
A	-		-		-
Morphoderivational Properties					

Inference Patterns

P34: This is an utter fool $\rightarrow$ This is a fool

Semantics

$\lambda N' \lambda Pol \lambda e \lambda x.\exists e_a \exists e_n \exists d[\textsf{N}_a(e_a) \wedge \textsf{ranking}(e_a, d) \wedge \textsf{role\_as}(e_a, n) \wedge Pol(\textsf{value}(d, adj)) \wedge \textsf{N'}(\lambda S.S)(e_n)(x)]$

(e.g. *utter*)

$\lambda N' \lambda Pol \lambda e \lambda x.\exists e_a \exists e_n \exists d[\textsf{evaluation}(e_a) \wedge \textsf{ranking}(e_a, d) \wedge \textsf{role\_as}(e_a, n) \wedge Pol(\textsf{value}(d, \textsf{utter})) \wedge \textsf{N'}(\lambda S.S)(e_n)(x)]$

A.2.19 Class Sa4

Members

first1, former1, intermediate1, last2, latter1, second1, third1

Features

Semantic Properties			
Subsectivity	Absoluteness	Transparency	
S	NA	NT	
Lexical Semantic Properties			
Gradability	Very-Mod	Antonymy	Stability
NG	-	M	D
Syntactic Properties			
Pred/Attr Construction	For/As Construction	Complementation	NP$_{meas}$
A	-	-	-
Morphoderivational Properties			

Inference Patterns

P34: This is the second chapter → This is a chapter

Semantics

$\lambda N' \lambda Pol \lambda e \lambda x.\exists e_a \exists e_n \exists d[\mathsf{N}_a(e_a) \wedge \textit{ranking}(e_a, d) \wedge \textit{role\_as}(e_a, n) \wedge Pol(\textit{value}(d, adj)) \wedge \mathsf{N'}(\lambda S.S)(e_n)(x)]$

(e.g. *second*)

$\lambda N' \lambda Pol \lambda e \lambda x.\exists e_a \exists e_n \exists d[\textit{numeral}(e_a) \wedge \textit{ranking}(e_a, d) \wedge \textit{role\_as}(e_a, n) \wedge Pol(\textit{value}(d, \textit{second})) \wedge \mathsf{N'}(\lambda S.S)(e_n)(x)]$

A.3 Privative Adjectives

A.3.1 Class PRpa1

Members

fabricated1, fabulous2, fanciful2, fictional2, fictitious1, fictitious2, fictive1, imaginary1, invented2, legendary2, mythical1, mythic1, nonexistent1, unreal1

Features

Semantic Properties			
Subsectivity	Absoluteness	Transparency	
PR	A	T	
Lexical Semantic Properties			
Gradability	Very-Mod	Antonymy	Stability
NG	-	B	S
Syntactic Properties			
Pred/Attr Construction	For/As Construction	Complementation	NP$_{meas}$
PA	-	SubjE(S-only)	-
Morphoderivational Properties			
N$_a$			

Inference Patterns

P1: This is a fictitious address $\leftrightarrow$ This address is fictitious
P3: X is a fictitious insurance provider $\leftrightarrow$ X is fictitious as an insurance provider
P4: It is fictitious to find them $\leftrightarrow$ To find them is fictitious
P10: This name is fictitious $\rightarrow$ This name is not real
P11: This name is not fictitious $\rightarrow$ This is a real name
P13: This is a fictitious hero $\rightarrow$ This is a fictitious entity
P14: This is a fictitious man and is a hero $\rightarrow$ This is a fictitious hero
P25: It is clear that this is a fictitious name $\leftrightarrow$ The fictitiousness of this name is clear
P31: This is a fictitious name $\rightarrow$ This is fictitious
P37: This is a fictitious name $\rightarrow$ This is not a name

Semantics

$\lambda N' \lambda Pol \lambda e \lambda x \exists e_a \exists e_n.[Pol(A(e_a)) \wedge x = e_a \wedge Pol(\neg \textsf{\textit{Exist}}(x)) \wedge N'(\lambda S.S)(e_n)(x)]$

(e.g. *fictitious*)

$\lambda N' \lambda Pol \lambda e \lambda x \exists e_a \exists e_n.[Pol(\textsf{\textit{fictitious}}(e_a)) \wedge x = e_a \wedge Pol(\neg \textsf{\textit{Exist}}(x)) \wedge N'(\lambda S.S)(e_n)(x)]$

A.3.2 Class PRpa2

Members

apparent2, artificial1, assumed1, bogus1, counterfeit1, fake1, false6,9, imitative3, inauthentic1, ostensible2, simulated1, spurious3

Features

Semantic Properties			
Subsectivity	Absoluteness	Transparency	
PR	NA	NT	
Lexical Semantic Properties			
Gradability	Very-Mod	Antonymy	Stability
NG	-	B	S
Syntactic Properties			
Pred/Attr Construction	For/As Construction	Complementation	NP_{meas}
PA	-	-	-
Morphoderivational Properties			
N_a, N_{cat}			

Inference Patterns

P1: This is a fake fur $\leftrightarrow$ This fur is fake
P10: This fur is fake $\rightarrow$ This fur is not genuine
P11: This fur is not fake $\rightarrow$ This fur is genuine
P14: This is a fake weapon and is a gun $\rightarrow$ This is a fake gun
P25: It is clear that this is a fake fur $\leftrightarrow$ The fakery of this fur is clear
P26: This is a fake fur $\leftrightarrow$ This fur is a fake
P37: This is a fake fur $\rightarrow$ This is not a fur

Semantics

$\lambda N' \lambda Pol \lambda x \exists e_a \exists e_n.[Pol(A(e_a)) \wedge \textit{role\_as}(e_a, n) \wedge Pol(N'(\lambda S.\neg S)(e_n)(x))]$

(eg. *fake*)

$\lambda N' \lambda Pol \lambda x \exists e_a \exists e_n.[Pol(\textit{fake}(e_a)) \wedge \textit{role\_as}(e_a, n) \wedge Pol(N'(\lambda S.\neg S)(e_n)(x))]$

A.3.3 Class PRpa3

Members

expected1, foreseen1, predicted1

Features

Semantic Properties			
Subsectivity	Absoluteness	Transparency	
PR	NA	NT	
Lexical Semantic Properties			
Gradability	Very-Mod	Antonymy	Stability
NG	-	B	D
Syntactic Properties			
Pred/Attr Construction	For/As Construction	Complementation	NP_{meas}
PA	as	SubjE(for)	-
Morphoderivational Properties			
V_{res}			

Inference Patterns

P1: This is a foreseen result ↔ This result is foreseen
P3: This is a foreseen result ↔ This is foreseen as a result
P4: It is foreseen to convert the area into a park ↔ To convert the area into a park is foreseen
P6: They are foreseen to be produced in the same foundry ↔ It is foreseen for them to be produced in the same foundry
P10: This result is foreseen → This result is not unexpected
P11: This result is not foreseen → This result is unexpected
P20: Someone has foreseen John as a candidate → John is a foreseen candidate
P37: John is the foreseen candidate → John is not a candidate

Semantics

$\lambda N' \lambda Pol \lambda e \lambda x \exists e_a \exists e_n.[Pol(V_a(e_a)) \wedge Pol(\textit{patient}(e_a, x)) \wedge \textit{state-of-affaires}(e_a, n)$
$\wedge N'(\lambda S. \neg S)(e_n)(x)]$

(e.g. *foreseen*)

$\lambda N' \lambda Pol \lambda e \lambda x \exists e_a \exists e_n.[Pol(\textit{foreseen}(e_a)) \wedge \textit{patient}(e_a, x) \wedge \textit{state-of-affaires}(e_a, n)$
$\wedge N'(\lambda S. \neg S)(e_n)(x)]$

A.3.4 Class PRpa4

Members

possible1, controversial1, impossible1, improbable1, likely1, probable2, uncertain1, unlikely1

Features

Semantic Properties			
Subsectivity	Absoluteness	Transparency	
PR	NA	NT	
Lexical Semantic Properties			
Gradability	Very-Mod	Antonymy	Stability
G	+	B	D
Syntactic Properties			
Pred/Attr Construction	For/As Construction	Complementation	NP$_{meas}$
PA	as	SubjE(S-only)	-
Morphoderivational Properties			
N$_a$			

Inference Patterns

P1: This is a possible solution $\leftrightarrow$ This solution is possible
P3: This is a possible solution $\leftrightarrow$ This is possible as a solution
P4: It is possible that it will rain tomorrow $\leftrightarrow$ That it will rain tomorrow is possible
P10: This result is possible $\rightarrow$ This result is not impossible
P11: This result is not possible $\rightarrow$ This result is impossible
P25: It is clear that this is a possible answer $\leftrightarrow$ The possibility of the answer is clear
P37: John is a possible candidate $\rightarrow$ John is not a candidate

Semantics

$\lambda N' \lambda Pol \lambda e \lambda x \exists e_a \exists e_n.[Pol(A'(e_a)) \wedge \textit{mod}(e_a, N) \wedge N'(\lambda S.\neg S)(e_n)(x)]$

(e.g. *possible*)

$\lambda N' \lambda Pol \lambda e \lambda x \exists e_a \exists e_n.[Pol(\textit{possible}(e_a)) \wedge \textit{mod}(e_a, N) \wedge N'(\lambda S.\neg S)(e_n)(x)]$

A.3.5 Class PRa1

Members

pretended1, pseudo1, would-be1, seeming1

Features

Semantic Properties			
Subsectivity	Absoluteness	Transparency	
PR	NA	NT	
Lexical Semantic Properties			
Gradability	Very-Mod	Antonymy	Stability
NG	-	B	S
Syntactic Properties			
Pred/Attr Construction	For/As Construction	Complementation	NP$_{meas}$
A	-	-	-
Morphoderivational Properties			
V$_a$			

Inference Patterns

P17: He is a pretended crime fighter ↔ He pretends to be a crime fighter
P37: He is a pretended esthete → He is not an esthete

Semantics

$\lambda Pol\lambda N\lambda x\exists e_a\exists e_n\exists e_{n1}.[V_a(e_a)\wedge \textit{agent}(e_a,x)\wedge \textit{state-of-affaires}(e_a,e_{n1})\wedge N(\lambda S.S)(e_{n1})(e_{n1})\wedge N(\lambda S.\neg S)(e_n)(x)]$

(e.g. *pretended*)

$\lambda Pol\lambda N\lambda x\exists e_a\exists e_n\exists e_{n1}.[\textit{pretend}(e_a)\wedge \textit{agent}(e_a,x)\wedge \textit{state-of-affaires}(e_a,e_{n1})\wedge N(\lambda S.S)(e_{n1})(e_{n1})\wedge N(\lambda S.\neg S)(e_n)(x)]$

A.3.6 Class PRa2

Members

eventual1, potential1

Features

Semantic Properties			
Subsectivity	Absoluteness	Transparency	
PR	NA	NT	
Lexical Semantic Properties			
Gradability	Very-Mod	Antonymy	Stability
NG	-	B	D
Syntactic Properties			
Pred/Attr Construction	For/As Construction	Complementation	NP$_{meas}$
A	-	ObjE	-
Morphoderivational Properties			
N$_a$			

Inference Patterns

P25: It is clear that John is a potential candidate $\leftrightarrow$ The potentiality of John as a candidate is clear
P37: John is a potential candidate $\rightarrow$ John is not a candidate

Semantics

$\lambda N' \lambda Pol \lambda e \lambda x \exists e_a \exists e_n.[Pol(A'(e_a)) \wedge \textit{mod}(e_a, N) \wedge N'(\lambda S.\neg S)(e_n)(x)]$

(e.g. *eventual*)

$\lambda N' \lambda Pol \lambda e \lambda x \exists e_a \exists e_n.[Pol(\textit{eventual}(e_a)) \wedge \textit{mod}(e_a, N) \wedge N'(\lambda S.\neg S)(e_n)(x)]$

A.3.7 Class PRa3

Members

early3, former2,3, future1, incoming1, last1, last2, late3, outgoing1, past1, preceding1, previous1, recent2, succeeding1

Features

<table>
<tr><td colspan="12">Semantic Properties</td></tr>
<tr><td colspan="4">Subsectivity</td><td colspan="4">Absoluteness</td><td colspan="4">Transparency</td></tr>
<tr><td colspan="4">PR</td><td colspan="4">NA</td><td colspan="4">NT</td></tr>
<tr><td colspan="12">Lexical Semantic Properties</td></tr>
<tr><td colspan="3">Gradability</td><td colspan="3">Very-Mod</td><td colspan="3">Antonymy</td><td colspan="3">Stability</td></tr>
<tr><td colspan="3">NG</td><td colspan="3">-</td><td colspan="3">M</td><td colspan="3">S</td></tr>
<tr><td colspan="12">Syntactic Properties</td></tr>
<tr><td colspan="3">Pred/Attr Construction</td><td colspan="3">For/As Construction</td><td colspan="3">Complementation</td><td colspan="3">NP$_{meas}$</td></tr>
<tr><td colspan="3">A</td><td colspan="3">-</td><td colspan="3">-</td><td colspan="3">-</td></tr>
<tr><td colspan="12">Morphoderivational Properties</td></tr>
<tr><td colspan="12">N$_{rel}$</td></tr>
</table>

Inference Patterns

P28: John is the former senator $\leftrightarrow$ John was senator in the past
P37: John is the former president $\rightarrow$ John is not the president

Semantics

$$\lambda N' \lambda Pol \lambda e \lambda x \exists e_a \exists e_n . Pol([A'(e_a) \wedge \mathit{time}(e_a, n) \wedge N'(\lambda S.S)(e_n)(x)])$$

(e.g. *former*)

$$\lambda N' \lambda Pol \lambda e \lambda x \exists e_a \exists e_n . Pol([\mathit{former}(e_a) \wedge \mathit{time}(e_a, n) \wedge N'(\lambda S.S)(e_n)(x)])$$

A.4 Plain Nonsubsective Adjectives

A.4.1 Class PlNS

Members

alleged2, purported1, putative1, reputed1, supposed3

Features

Semantic Properties			
Subsectivity	Absoluteness	Transparency	
NS	NA	NT	
Lexical Semantic Properties			
Gradability	Very-Mod	Antonymy	Stability
NG	-	B	D
Syntactic Properties			
Pred/Attr Construction	For/As Construction	Complementation	NP_{meas}
A	-	SubjE(of)	-
Morphoderivational Properties			
V_{res}			

Inference Patterns

P5: He is alleged to have unintentionally murdered his classmate $\leftrightarrow$ It is alleged of him that he had unintentionally murdered his classmate

P20: Someone thinks John is a murderer $\rightarrow$ John is an alleged murderer

Semantics

$\lambda N' \lambda Pol \lambda e \lambda x \exists e_a \exists e_n.[Pol(A'(e_a, e_n)) \wedge x = e_a \wedge N'(\lambda S.S)(e_n)(e_n)]$

(e.g. *alleged*)

$\lambda N' \lambda Pol \lambda e \lambda x \exists e_a \exists e_n.[alleged(e_a, e_n) \wedge x = e_a \wedge N'(\lambda S.S)(e_n)(e_n)]$

or

$\lambda N' \lambda Pol \lambda e \lambda x \exists e_a \exists e_n.[\textit{say}(e_a) \wedge \textit{patient}(e_a, x) \wedge \textit{topic}(e_a, e_n) \wedge N'(\lambda S.S)(e_n)(e_n)]$

Bibliography

Joannes M.G.A. Aarts. Adjective-noun combinations: A model for their semantic interpretation. 1976.

S. Ait-Mokhtar, J P. Chanod, and C. Roux. Robustness beyond shallowness: incremental deep parsing. *Natural Language Engineering*, 8(2/3):121–144, 2002.

Marilisa Amoia and Claire Gardent. Recognition of alternation paraphrases: a robust and exhaustive approach. In *Proceedings of KRAQ'05 (Knowledge and Reasoning for Answering Questions),* Edinburgh, Scotland, 2005.

Marilisa Amoia and Claire Gardent. A test suite for inference involving adjectives. In *Proceedings of the 6th Language Resources and Evaluation Conference (LREC)*, Marrakech, Morocco, 2008.

R. Stephen Anderson. On the role of deep structure in semantic interpretation. *Foundations of Language*, 7(3):387–396, 1971.

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

Collin F. Baker, Charles J. Fillmore, and John B. Lowe. The berkeley framenet project. In *In Proceedings of COLING-ACL*, Montreal, Canada, 1998.

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

Satanjeev Banerjee and Ted Pedersen. Extended gloss overlaps as a measure of semantic relatedness. In *Proceedings of the 18th International Joint Conference on Artificial Inteligence,* Acapulco, Mexico, 2003.

Roy Bar-Haim, Ido Dagan, Bill Dolan, Lisa Ferro, Danilo Giampiccolo, Bernardo Magnini, and Idan Szpektor. *The Second PASCAL Recognising Textual Entailment Challenge*. 2006.

R. Bartsch and T. Vennemann. *The Comparative*. Athäneum Verlag, Frankfurt, 1973.

Regina Barzilay and L. Lee. Learning to paraphrase: an unsupervised approach using multiple-sequence alignment. In *Proceedings of NAACL-HLT,* 2003.

Regina Barzilay and Kathleen McKeown. Extracting paraphrases from a parallel corpus. In *Proceedings of ACL/EACL*, 2001.

John A. Bateman. Upper modeling: A level of semantics for natural language processing. In *Proceedings of the Fifth International Natural Language Generation Workshop*, pages 54–60, 1990.

John A. Bateman. The merged upper model: a linguistic ontology for german and english. In *Proceedings of the 15th conference on Computational Linguistics*, pages 803–809, 1994.

S. Beale, S. Nirenburg, and K. Mahesh. Semantic analysis in the mikrokosmos machine translation project. In *In Proceedings of the 2nd Symposium on Natural Language Processing*, pages 297–307, Bangkok, Thailand, 1995.

M. Bierwisch. *The Semantics of Gradation*. Springer Verlag, Berlin, 1989.

Patrick Blackburn and Johan Bos. *Representation and Inference for Natiral Language. A first Course in Computational Semantics*. CSLI Studies in Computational Linguistics, 2005.

Johan Bos. Towards wide-coverage semantic interpretation. In *In Proceedings of the Sixth International Workshop on Computational Semantics IWCS-6*, pages 42–53, 2005.

C. Brew and S. Schulte im Walde. Spectral clustering for german verbs. In *Conference on Empirical Methods in Natural Language Processing*, Philadelphia, USA, 2002.

E. J. Briscoe and J. Carroll. Automatic extraction of subcategorization from corpora. In *In Porceedings of the 5th ACL Conference on Applied Natural Language Processing*, Washington DC, USA, 1997.

C. Brun and C. Hagège. Normalization and paraphrasing using symbolic methods. In *ACL: Second International workshop on Paraphrasing, Paraphrase Acquisition and Applications*, pages 41–48, Sapporo, Japan, 2003.

C. Quirk C., C. Brockett, and W. B. Dolan. *Monolingual Machine Translation for Paraphrase Generation*. Barcelona, Spain, 2004.

G.N. Carlson. A unified analysis of the english bare plural. *Linguistics and Philosophy*, pages 413–455, 1977.

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

Noam Chomsky. *Syntactic Structures*. Mouton, The Hague, 1957.

Noam Chomsky. *Current Issues in Linguistic Theory*. The Hague, The Netherlands: Mouton de Gruyer, 1964.

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

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

Hoa Trang Dang, Karin Kipper, Martha Palmer, and Joseph Rosenzweig. Investigating regular sense extensions based on intersective levin classes. In *Proceedings of COLING-ACL98*, Montreal, Canada, August 1998.

Hoa Trang Dang, Karin Kipper, and Martha Palmer. Integrating compositional semantics into a verb lexicon. In *COLING-2000 Eighteenth International Conference on Computational Linguistics*, Saarbruecken, Germany, July 31-August 4 2000.

J. Deese. The associative structure of some english adjectives. *Journal of Verbal Learning and Verbal Behaviour*, 3:347–357, 1964.

J. Deese. *The structure of associations in language and thought*. Johns Hopkins Press, Baltimore, 1965.

R.M.W. Dixon. *Where Have All the Adjectives Gone?* Berlin-Amsterdam-New York:Mouton, 1982.

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

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

B. Dorr. Large-scale dictionary construction for foreign language tutoring and interlingual machine translation. *Machine Translation*, 12(4):271–325, 1997.

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

C. Fellbaum. English verbs as a semantic net. *International Journal of Lexicography*, 3 (4):278–301, 1990.

C.J. Fillmore. Frame semantics and the nature of language. In *In Annals of the New York Academy of Sciences: Conference on the Origin and Development of Language and Speech*, volume 280, pages 20–32, 1976.

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

William Frawley. *Linguistic Semantics*. Erlbaum, Hillsdale, N.J., 1992.

Atsushi Fujita. *Automatic Generation of Syntactically Well-formed and Semantically Appropriate Paraphrases*. PhD Dissertation, Nara Institute of Science and Technology, 2005.

O. Glickman and I. Dagan. Identifying lexical paraphrases from a single corpus: a case study for verbs. In *Proceedings of Recent Advances in Natural Language Processing*, 2003.

Ralph Grishman, Catherine Macleod, and Adam Meyers. Comlex syntax: Building a computational lexicon. In *In Proceedings of the 15th International Conference on Computational Linguistics (COLING 94)*, Kyoto, Japan, August 1994.

D. Gross and Katherine J. Miller. Adjectives in wordnet. *International Journal of Lexicography*, 3(4):265–277, 1990.

XTAG Research Group. *A lexicalized tree adjoining grammar for english*. Technical Report IRCS-01-03, University of Pennsylvania, 2001.

C. Hagège and C. Roux. Entre syntaxe et sémantique : normalisation de la sortie de l'analyse syntaxique en vue de l'amèlioration de l'extraction d'information à partir de textes. In *TALN*, Batz-sur-Mer,France, 2003.

Z. Harris. *Distributional Structure*. New York:Oxford University Press, 1985.

Z. Harris. *Methods in Structural Linguistics*. Chicago: University of Chicago Press, 1951.

I. Heim. Notes on comparatives and related matters. *Ms. University of Texas*, 1985.

A. Hickl, J. Williams, J. Bensley, K. Roberts, B. Rink, and Y. Shi. Recignizing textual entailment with lcc's groundhog system. In *In Proceedings of the PASCAL RTE-2 Challenge*, 2006.

Jerry R. Hobbs. Ontological promiscuity. In *Proceedings of the 23rd Annual Meeting of the Association for Computational Linguistics*, pages 61–69, Chicago, Illinois, July 1985.

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

Ray Jackendoff. *Semantic structures*. The MIT Press, 1990.

M. Kaisser and B. Webber. Question answering based on semantic roles. In *In Proceedings of the ACL 2007 Deep Linguistic Processing Workshop, ACL-DLP 2007*, Prague, Czech Republic, 2007.

Hans Kamp. Two theories about adjectives. In Edward L. Keenan (ed.), *Formal Semantics of Natural Language*, pages 123–155. Cambridge University Press, 1975.

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

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

Edward Keenan and Leonard Faltz. Boolean semantics for natural language. *Synthese Language Library*, (23), 1985.

C. Kennedy and L. McNally. Scale structure and the semantic typology of gradable predicates. *Language*, 81(2):1–37, 2005.

Christofer Kennedy. *Comparatives, Semantics of*. Keith Allen (section editor), Elsevier, Oxford, 2004.

Christofer Kennedy. Vagueness and grammar: The semantics of relative and absolute gradable adjectives. *Ms.*, pages 129–191, June 2005.

Christofer Kennedy and Louise McNally. *From Event Structure to Scale Structure: Degree Modification in Deverbal Adjectives*. Ithaca, 1999.

Paul Kingsbury, Martha Palmer, and Mitch Marcus. Adding semantic annotations to the penn treebank. In *Proceedings of the Human Language Technology Conference*, San Diego, California, 2002.

Karin Kipper, Hoa Trang Dang, and Martha Palmer. Class-based construction of a verb lexicon. In *Proceedings of the 7th National Conference on Artificial Intelligence (AAAI-2000)*, Austin, TX, July-August 2000. Morgan Kaufmann.

K. Kipper-Schuler. *VerbNet: A broad-coverage, comprehensive verb lexicon, PhD dissertation*. University of Pennsylvania, 2005.

J. L. Klavans and M. Kan. Role of verbs in document analysis. In *Proceedings of COLING/ACL*, pages 680–686, Montreal, Canada, 1998.

E. Klein. *Comparatives*. Walter de Gruyter, Berlin, 1991.

A. Korhonen and E. J. Briscoe. Extended lexical-semantic classification of english verbs. In *In Porceedings of the HLT/NAACL Workshop on Computational Lexical Semantics*, Boston, MA, 2004.

Angelika Kratzer. *Stage-Level and Individual-Level Predicates*. G.N. Carlson and F.J. Pelletier (eds.), The generic book, Chicago and Londres: Chicago University Press, 1995.

M. Lapata. A corpus-based account of regular polysemy: The case of context-sensitive adjectives. In *In Proceedings of the 2nd Meeting of the NAACL*, Pittsburgh, PA, 2001.

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

Beth Levin. *English Verb Classes and Alternation, A preliminary Investigation*. The University of Chicago Press, Chicago, 1993.

Dekang Lin and Patrick Pantel. Discovery of inference rules for question answering. *Natural Language Engineering*, 2001.

J. Lyons. *Semantics*. Cambridge University Press, New York, 1977.

I. Melcuk. Paraphrase et lexique dans la theorie linguistique sens-texte. *Lexique*, 6: 13–54, 1988.

I. Melcuk and A. Zolkovskij. Towards a functioning 'meaning-text' model of language. *Linguistics*, 57:10–47, 1970.

I. Melcuk and A. Zolkovskij. Explanatory combinatorial dictionary of modern russian. 1984.

P. Merlo and S. Stevenson. Automatic verb classification based on statistical distributions of argument structure. *Computational Linguistics*, 27(3):373–408, 2001.

G.A. Miller, R. Beckwith, C. Fellbaum, D.Gross, and K. Miller. Introduction to wordnet: An on-line lexical database. In *Communications of the ACM 38 (11)*, 1993.

K. J. Miller. Modifiers in wordnet. In C. Fellbaum (ed.), *WordNet An Electronic Lexical Database*. Cambridge, MA, The MIT Press, 1998.

Richard Montague. *The proper treatment of quantification in ordinary English*. In Hintikka, K.J.J., Moravcsik, J.M.E. and Suppes, P. (eds.), Approaches Natural Language, Dordrecht: Reidel, 1973.

S. Nirenburg, J. Carbonell, M. Tomita, and K. Goodman. *Machine Translation: A Knowledge-Based Approach*. Morgan Kaufmann, San Mateo, CA, 1992.

Stephan Oepen and Klaus Netter. Tsnlp - test suites for natural language processing. In *Conference on Linguistic Databases*, Groningen, The Netherlands, 1995.

Terence Parson. Thematic relations and arguments. *Linguistic Inquiry*, 26(4):635–662, 1995.

I. Peters. Simple english lexicon report. Technical report, March, 2000.

I. Peters and W. Peters. *The Treatment of Adjectives in SIMPLE: Theoretical Observations*. Athens, 2000.

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

V. Raskin and S. Nirenburg. *Lexical Semantics of Adjectives, a micro-theory of adjectival meaning*. MCCS Report, 1995.

Deepak Ravichandran and Eduard Hovy. Learning surface text patterns for a question answering system. In *Proceedings of 40th Annual Meeting of the Assoc. of Computational Linguistics (ACL)*, pages 215–222, 2002.

Josef Ruppenhofer, Michael Ellsworth, Miriam R. L. Petruck, Christopher R. Johnson, and Jan Scheffczyk. *FrameNetII: Extended Theory and Practice*. 2006.

Neville Ryant and Karin Kipper. Assigning xtag trees to verbnet. In *Proceedings of the 7th International Workshop on Tree Adjoining Grammar and Related Formalisms (TAG+7)K*, Vancouver, British Columbia, Canada, 2004.

Roger Schwarzschild and Karina Wilkinson. Quantifiers in comparatives: A semantics of degree based on intervals. *Natural Language Semantics*, 10:1–41, 2002.

P.A. Seuren. *The Comparative*. Dordrecht: Riedel, 1973.

Y. Shinyama, S. Sekine, K. Sudo, and R. Grishman. Automatic paraphrase acquisition from news articles. In *Proceedings of HLT*, 2002.

M. E. A. Siegel. Capturing the adjective. 1976.

G. Silva and S.A. Thompson. On the syntax of adjectives with 'it' subject and infinitival complements in english. *Studies in Language*, 1(1):109–126, 1977.

Specification Group SIMPLE. Specification simple work package 2. Linguistic specifications deliverable d2.1, 2000.

M. Stede. A generative perspective on verb alternations. *Computational Linguistics*, 24 (3):401–430, 1998.

Koichi Takeuchi. Developing an lcs dictionary for verbs. In *Proceedings of the 10th Annual Meeting of the Association for Natural Language Processing*, pages 576–579, 2004.

M. Tatu, B. Iles, J. Slavik, A. Novischi, and D. Moldovan. Cogex at the second recognizing textual entailment challenge. In *In Proceedings of the PASCAL RTE-2 Challenge*, 2006.

TREC. *The Sixteenth Text REtrieval Conference Proceedings (TREC 2007)*. Department of Commerce, National Institute of Standards and Technology, Gaithersburg, Maryland, 2007.

TREC-8. *The Eighth REtrieval Conference (TREC-8)*. Department of Commerce, National Institute of Standards and Technology, Gaithersburg, Maryland, 1999.

F. Trouilleux. These de doctorat en sciences du langage. *Universite Blaise Pascal, Clermont-Ferrand, France*, 2001.

Hans Uszkoreit and Stanley Peters. On some formal properties of meta-rules. *Linguistics and Philosophy*, 9(4):477–494, 1986.

L. Venderwende, D. Coughlin, and B. Dolan. What syntax can contribute in entailment task. In *In Proceedings of the PASCAL Wokshop on Recognising Textual Entailment*, 2005.

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

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

A. von Stechow. Comparing semantic theories of comparison. *Journal of Semantics*, 3: 1–77, 1984.

A. Zaenan, L. Karttunen, and R. Crouch. Local textual inference: Can it be defined or circumscribed? In *In Proceedings of the ACL Workshop on Empirical Modeling of Semantic Equivalence and Entilment*, Ann Arbor, Michigan, 2005.

Index of English Adjectives