The Description of Adjectives for Natural Language Processing: Theoretical and Applied Perspectives

Pierrette Bouillon and Evelyne Viegas

ISSCO/ETI (Université de Genève) 54 route des ACACIAS 1227 Genève Suisse
New Mexico State University Computing Research Laboratory Las Cruces, NM 88003, USA

Abstract

Adjectives constitute a challenging issue for NLP applications. ¿From a syntactic viewpoint, they can be predicative or attributive. From a semantic viewpoint, their sense can vary depending on the context in which they appear. In this paper we present an overview on the lexical semantics of adjectives in Natural Language Processing, followed by the presentation of the papers included in this proceeding.[1]

1. Introduction

Adjectives constitute a challenging issue for NLP applications. From a syntactic viewpoint, they can be predicative or attributive. From a semantic viewpoint, their sense can vary depending on the context in which they appear. For instance, *difficult* does not have the same meaning in the phrases *a difficult book*, *a difficult child* and *a difficult exam*. In French "vieux" (old) does not have the same meaning in the phrases "un homme vieux" (an old/aged man) and "un vieux maire" (a mayor of long standing, not necessarily old). No NLP system can ignore adjectives: in order to correctly analyze, generate, translate natural languages or index a text, the system must understand (or at the least, simulate) the meaning of the adjective.

In the proposed workshop we have brought together linguists and computational linguists to discuss theoretical and applied research on the treatment of adjectives in the various fields of NLP. Burning issues include:

lexical semantics of adjectives

types of treatments: enumerative, generative, contextual

analysis, generation, translation and indexing of adjectives

compositional semantics of adjectives

semi-compositional semantics of adjectives (collocations in Adj-Noun constructions)

---

[1] Part of this document has been printed for EAGLES, Lexicon Interest Group on semantic encoding relative to the period May through December 1997.

In the following, we first present an account of the lexical semantics of adjectives in NLP (in terms of their basic properties and of how they are represented in lexical semantic theories). Then we look at adjectives as part of co-occurrences (i.e. how to predict their restrictions at the syntagmatic level). Finally, we briefly introduce the papers presented at the workshop.

2. Lexical Semantics of Adjectives

Adjectives have not much been studied in traditional lexical semantics compared to the large amount of work devoted to verbs and to nouns. They nevertheless present an interesting polymorphic behavior. Syntactically, they can appear in different positions in the sentence, as modifier of the noun (*a sad book*) or complement of a copular verb like *be* (*this book is sad*). Semantically, adjectives, more than other categories, are able to take different meanings depending on their context (for example *difficult* in *a difficult child*, *a difficult book*, *a difficult exam* or *fast* in *a fast car*, *a fast motorway*, *a fast procedure*, etc.) ((Marx, 1993), (Lahav, 1989), etc.). As senses are only restricted by the number of possible contexts, it is questionable if they can be enumerated. Even more, richer representations and compositional mechanisms seem to be required (Bouillon, 1997).

In the following, we give first a general overview of this polymorphism. We then examine how they are represented in three models: relational (*WordNet*, §2.2.2), generative (Generative Lexicon, §2.2.3) and ontological (MikroKosmos, §2.2.4).

2.1. Classification of adjectival polymorphism

Adjectives differ in many ways. We will examine their polymorphism from three different perspectives: syntactic, semantic and ontological.

2.1.1. Syntactic classes

Syntactically, adjectives can be classified with respect to three features: function, complementation and alternation.

1. **Function:** adjectives can appear in *attributive* position, as noun modifiers inside a Noun Phrase (NP) as shown in (1), or in *predicative* position as a complement of a verb like *be, seem, consider*, etc., as shown in (2).

    (1)   A happy person

    (2)   They are happy, they consider him happy, he seems happy, etc.

    This criteria allows us to distinguish three different types of adjectives: *predicative-only* (3) and *attributive-only* (4) for adjectives that are only used in one position, or *central* for those used both predicatively and attributively, as *tall* in (5) ((Quirk *et al.*, 1994), for example).

    (3)   *Afraid people, people are afraid

    (4)   The atomic scientist, *the scientist is atomic

    (5)   The tall man, the man is tall

When attributive, adjectives can be both *postnominal* and *prenominal* (6a), even if the postnominal position is less common in English (Sadler and Arnold, 1993). But there are restrictions: an adjective like *former* cannot be postnominal (6b) and some adjectives appear only after the noun (6c).

(6) a. Navigable rivers, rivers navigable

b. Former architect, *architect former

c. *Aforethought malice, malice aforethought

2. **Complementation:** As verbs, adjectives differ in terms of their argument structure. Many adjectives accept no complement at all (for example, *Belgian*, *red*, etc.). Those that accept complements can be subclassified as follows:

- the type of the complement they subcategorize, i.e. prepositional as in (7a), or clausal (verb phrase as in (7b) or sentence as in (7c));

(7) a. I'm proud of you

b. I'm sad to leave

c. It is possible that they leave

- whether the complement is optional or not (*I'm desirous to leave, *I'm desirous* vs *I'm sad to leave, I'm sad*);

- for prepositional complements, the preposition that marks the NP (for example *capable of, skillful at, absent from*, etc.).

3. **Alternations:** As for verbs, adjectives enter into alternations.

- Among adjectives that take clausal complement, two subclasses can be distinguished according to whether this complement may or may not be realized as the subject. It is common to call *object embedding* adjectives whose complements cannot be realized as subject (8) and *subject embedding* those which can take a clausal subject (9) (see for example (Picabia, 1978), (Vendler, 1968), (Arnold, 1989), (Dixon, 1991) and (Silva and Thompson, 1977)).

(8) a. i. I'm sad to leave
        ii. To leave is sad for me

b. i. They are certain to leave
        ii. To leave is certain for them

c. i. Sam was brave to leave
        ii. To leave was brave of Sam

(9) a. i. I'm eager to come
        ii. *To come is eager

The adjectives that allow their subject to be clausal can be further classified depending on what preposition marks the accompanying nominal complement, as in (10) ((Arnold, 1989); (Silva and Thompson, 1977)). This allows (Arnold, 1989) to distinguish four classes of adjectives: *S-only* and *S+toNP* adjectives indicate some

perception of the modality of the clausal clause; *S+ofNP* adjectives give an evaluation of an individual, on the presupposition that he is responsible for the state of affairs described in the proposition. Finally, *S+forNP* adjectives characterize directly a state of affairs, but may also indicate an experiencer.

(10) a. That they left is possible (S-only)

b. That they left is clear to Sam (S+toNP)

c. That he left was brave of him (S+ofNP)

d. To ignore pollution is dangerous for the country (S+forNP)

- *It-extraposition*: All adjectives that are subject-embedding allow the it-extraposition (11).

(11) a. It is possible that they left

b. It is clear to Sam that they left

c. It was brave of him that he left

d. It is dangerous for the country to ignore pollution

- *Tough-construction*: some *s+forNP* adjectives allow the tough-construction (12b), where the subject of the adjective is understood as filling a non-subject position in the complement of the adjective (see for example (Arnold, 1989) for an overview).

(12) a. It is difficult to support this proposal

b. This proposal is difficult to support

2.1.2. *Logical classes*

Semantically, adjectives can belong to three different classes, which differ in their logical behavior in the following way ((Chierchia and McConnell-Ginet, 1990); (Arnold, 1989); (Parsons, 1990), pp. 43-44).

1. An adjectives (ADJ) is said *absolute* (or *intersective*, *predicative*, etc.) if (13a) implies (13b) and (13c). These adjectives are standardly analyzed as predicates: they denote properties and the denotation of the adjective-noun construction is the intersection of the denotation of the ADJ and the N.

(13) a. this is a red (ADJ) table (N)

b. → this individual is a N

c. → this individual is ADJ

Typical examples of this category are adjectives which denote:

i. a shape: *hexagonal*

ii. a social group or a nationality: *communist, Belgian*, etc.

iii. a color

2. An adjective is *property-modifying* (or *non-intersective*, *operators*, etc.) if (14a) does not imply (14b), nor often (14c): *a former architect* is not *an architect*, nor *former*. These adjectives have been analyzed as operators: they denote a function from properties to properties. Thus, for example, *former* in (14a) takes the denotation of *architect* to the set of individuals who used to be architect.

    (14)   a.   this is a former (ADJ) architect (N)

                 b. $/\rightarrow$ this individual is a N

                 c. $/\rightarrow$ this individual is ADJ

Property-modifying adjectives include: nominal (or relational) adjectives (*polar bear*, *atomic scientist*, etc. Cf. (Levi, 1978)), manner (or adverbial) adjectives (*a poor liar*, *a fast car*), emotive (*a poor man*) and modals, i.e. all adjectives which are related to adverbs, quantifiers or determiners (*a feeble excuse*, *the specific reason*, *a fake nose*, etc.).

3. An adjective is *relative* (or *scalar*) if (15a) implies (15b), but not (15c); *a big mouse*, for example, is not *a big animal*. As absolute adjectives, they characterize the individual described by the noun (15b), but, unlike them, it is relative to some norm or standard of comparison: *a big mouse* is *big for an F*, where F is supplied by the context (15d). As they share properties with absolute and property modifying adjectives, they have been analyzed both as predicates and operators ((Parsons, 1990), p. 44): on the predicative treatment, *x is a clever N* for example, means therefore *x is an N & x is clever for an F* and on the operator treatment, it means *clever(x is an N that is F)*.

    (15)   a.   this is a big mouse

                 b. $\rightarrow$ this individual is a N

                 c. $/\rightarrow$ this individual is Adj

                 d. $\rightarrow$ this individual is Adj for a F

2.1.3. *Other semantic classes*

The adjectives can also be classified with respect to other semantic features (Quirk *et al.*, 1994), as:

1. **Aspect:** an adjective can be *stage-level* (if it expresses a temporary or accidental property) as in (16) or *individual-level* (in case of a generic, permanent or inherent property), see (17) ((Carlson, 1977); (Carlson and Pelletier, 1995)).

    (16)   *drunk, available* , etc.

    (17)   *clever, tall*, etc.

2. **Gradation:** an adjective can be gradable or not.

2.1.4. Adjective taxonomies

Adjective taxonomies classify adjectives in the different semantic categories they can express (see (Raskin and Nirenburg, 1995) for a good introduction). (Dixon, 1991) is one of the most representative. He classifies adjectives as the following:

(18) 1. DIMENSION: *big, short*, etc.

2. PHYSICAL PROPERTY: *strong, ill*, etc.

3. SPEED: *fast, quick*, etc.

4. AGE: *new, old*, etc.

5. COLOR: *red, black*, etc.

6. VALUE: *good, bad*, etc.

7. DIFFICULTY: *easy, difficult*, etc.

8. QUALIFICATION: DEFINITE (*probable*), POSSIBLE (*possible*), USUAL (*usual*), LIKELY (*likely*), SURE (*sure*), CORRECT (*appropriate*)

9. HUMAN PROPENSITY: FOND (*fond*), ANGRY (*angry*, *jealous*), HAPPY (*anxious*, *happy*), UNSURE (*certain*), EAGER (*eager*, *ready*), CLEVER (*clever*, *stupid*, *generous*)

10. SIMILARITY: *similar*, *different*, etc.

For each class, Dixon specifies the syntactic behavior of each adjective, as follows for the EAGER class: "EAGER takes an NP or a THAT or MODAL (FOR) TO complement, e.g. *I'm eager for the fray*, *I'm eager that Mary should go*, *I'm eager (for Mary) to go*. *Ready* may only take an NP or a Modal (FOR) TO clause (not a THAT complement) while *willing* must take a THAT or Modal (FOR) TO clause, i.e. it cannot be followed by proposition plus NP." ((Dixon, 1991), p. 83).

These classifications have a descriptive and a potential heuristic value (each distinction could suggest the introduction of an ontological feature), but none of these types of taxonomies have been applied in practical applications ((Raskin and Nirenburg, 1995), p. 10).

2.2. Representation of the properties

2.2.1. Introduction

The previous section identified the different properties of adjectives. This one examines how they are represented and generalized in three major models: *WordNet*, *Generative Lexicon* and *MikroKosmos*.

2.2.2. Semantic network, WordNet

In *WordNet* (§2.2.2), adjectives are divided in two main classes which are said to account for the majority of adjectives: *ascriptive* which are considered to ascribe a value of an attribute to a noun and *nonascriptive* which are similar to nouns used as modifiers (Gross and Miller, 1990).

Ascriptive adjectives are organized in terms of antonymy and synonymy, as in (19):

(19)   {DRY, WET1,! anhydrous,& arid,& …}
       {anhydrous, dry1, & }
       {arid, waterless,dry1, & }
       {dehydrated, desiccated, parched, dry1, & }
       {dried, dry1, & }
       …

Non-ascriptive ones are considered as stylistic variants of modifying nouns and are cross-referenced to the noun files. For example, the entry (20) indicates that *astral* and *stellar* have the meaning of *pertaining to a star or stars.*

(20)   {star | astral, stellar}

Gradation is not indicated in *WordNet* because it is not often lexicalized in English. Restrictions on syntactic position (for prenominal-only and postnominal-only adjectives) are encoded for specific words, as they cannot be generalized over entire sets of synonyms (21).

(21)   AWAKE(p), ASPEEP(p),! alert, & ALERT, …

WordNet does not say anything about the way senses are related: adjectives have as much senses as synsets. Moreover, it does not provide the means to predict grammatical properties from the representation (complementation, alternations, selective restriction). These two features distinguish the relational approach of *WordNet* from a Generative approach, like Generative Lexicon.

2.2.3.   Generative Lexicon

The Generative Lexicon (Pustejovky, 1995) focuses on the two aspects neglected in *WordNet*:

1. how the different adjectival senses are related and how they can be derived compositionally from the representations of the noun and the adjective (Bouillon, 1997), and

2. the syntax-semantics interface. In this theory, the adjectival polymorphism is explained by richer representations of adjectives and nouns (the qualia structure) and the way they combine together.

Take as an example the ambiguity of the French adjective "vieux" (old) in "un vieux marin" (an old sailor) which can be both relative (aged) and property modifying (with the meaning *who has this capacity for a long time, without being necessarily aged*). It can be explained in the following way (cf. (Bouillon, 1997) and (Bouillon, 1998) for more details).

1. Semantics of the noun: nouns to which adjectives apply have complex representations, which define the different predicates necessary for the definition of the word (Busa, 1996); a sailor, for example, is defined as *somebody who has the capacity of sailing*, i.e. as the conjunction of three types: human, state (*have the capacity to*) and event (*to sail*).

2. Semantics of the adjective: the semantics of "vieux" indicates that this adjective has two functional types: it can apply to an individual (" un vieil homme") or a state ("une vieille amitié").

3. Composition: As "vieux" has two different functional types and as "marin" is defined by the conjunction of these types, it can apply to both of them, giving rise to the ambiguity of the adjective-noun construction. In one sense, the adjective applies directly to the type *human* denoted by "marin" ("un vieux marin" is relative and is understood as *an old individual who has the capacity of sailing*); in the other, it subselects a part of the denotation of the noun, i.e the state *have a capacity* (" un vieux marin" is Property Modifying (PM) and is then interpreted as *somebody who has had this capacity for a long time*).

4. Syntax: Depending on the type the adjective applies to, the adjective enters into different relations with the noun: when the adjective applies to the head, it can form a property independent from the noun and be *intersective* (*relative*); when it modifies another type, it is necessarily in an *inclusion* relation: it forms one single property with the noun it modifies (it is PM). In French, the relation seems to be partly reflected at the syntactic level: (Bouillon, 1997) showed that an adjective can apparently appear both postnominally and prenominally if the composition is not ambiguous, i.e. if the adjective can only apply to one type. The prenominal position is however required to force a different relation from the postnominal one, in particular it allows the adjective to modify another type than the head of the qualia structure (i.e. the basic denotation of the noun). para.

With this kind of treatment, adjectives which belong to different logical classes are not considered as homonymous. The different senses can be derived from the semantics of the noun and the adjective.

2.2.4. Ontological approach, MikroKosmos

In a ontological approach like *MikroKosmos*, the lexicon mediates between the interlingual representation and the ontology. Each lexical entry indicates both the lexical mapping from language units to ontological concepts (in the LEX-MAP zone) and the linking between the syntactic and the semantic structures (SYN-STRUC). In *MikroKosmos*, the lexical entry needs not to be similar to the concept and can add specific constraints to it, as we will see for adjectives.

At the semantic level, the classification is more complete than in *WordNet*. The theory distinguishes adjectives which modify a noun both syntactically and semantically and those which can only modify the noun syntactically (cf. (Raskin and Nirenburg, 1995) and (Raskin and Nirenburg, 1996)).

For the former ones, the meaning is associated with a region on a scale which is defined as the range of an ontological property. The contribution that the adjective makes to the semantics of the sentence consists of inserting its meaning (a property-value pair) as a slot filler in a frame representing the meaning of the noun which this adjective modifies. The AN construction *big house* for example is represented as follows:

(22)   (private-home
         (size-attribute (value > 0.75)).

The latter ones, which cannot be reduced to the standard property-based type of adjectival modification, include: adjectives that denote attitudes (as *good*), and temporal, membership (*fake*), event-related (*abusive*) and denominal adjectives (*medical*). They are represented as follows:

1. **Attitude:** The meaning of these adjectives introduces an attitude on the part of the speaker with regard to the modified noun, as *good* in the following:

   (23)   (LEX-MAP
          (attitude
          (type evaluative
          (attitude-value (value > 0.75))
          (relaxable-to (value > 0.6))))).

2. **Temporal:** as temporal knowledge is represented with the meaning of the entire proposition, the adjective entries are not marked for it.

3. **Membership:** they focus on two elements: first, whether the modified noun is a member of a certain set, and second, whether the properties of this noun intersect significantly with those of the set members. For example, the property of *fake* members overlap with those of the set members only on unimportant properties, as physical resemblance.

4. **Event-related:** their semantics is derived from the verbal entry. Their meaning depends on the thematic role filled by the noun modified by the adjective.

5. **Denominal:** these adjectives are object-related. Their semantics specifies the exact relation they have with the noun, for example *medical* PERTAIN-TO *medicine*, *federal* is OWNED-BY a federation, *malignant* HAS-AS-PART cancer cells, etc.

At the syntactic level, the theory indicates the different subcategorizations of the adjective (as modifiers and/or predicate).

Following (Raskin and Nirenburg, 1996), the lexicon contains 6,000 English and 1,500 Spanish adjectives.

3.   Co-occurrences and Adjectives in NLP

In this section we look at adjectives as part of co-occurrences. We address two approaches which have been modeling co-occurrences and adjectives for NLP purposes using a linguistic approach à la Mel'čuk (Wanner, 1996) and a conceptual approach ((Viegas, 1998); (Nirenburg and Nirenburg, 1988)).

3.0.5.   *The Meaning Text Theory Approach*

Mel'čuk's Meaning Text Theory (MTT) is a generator-oriented lexical grammatical formalism. Lexical knowledge is encoded in an entry of the Explanatory Combinatorial Dictionary (ECD), each entry being divided into three zones: the *semantic zone* (a semantic network representing the meaning of the entry in terms of more primitive words), the *syntactic zone* (the grammatical properties of the entry) and the *lexical combinatorics zone* (containing the values of the **Lexical Functions** (LFs) [2]), central to collocations:

> *A lexical function F is a correspondence which associates a lexical item L, called the key word of F, with a set of lexical items F(L)-the value of F. (Mel'čuk, 1988).*

---

[2]The reader can consult (Iordanskaja *et al.*, 1991) and (Ramos *et al.*, 1995) for the use of LFs in MTT and NLG respectively.

We focus here on *syntagmatic* LFs describing co-occurrence relations such as *heavy smoker; legitimate complaint*.[3] Various LFs are used depending on the lexical semantics relationship between the adjective and the noun, as examplified below.

The adjective is used in its "standard" sense and is in no way figurative: such as in *rancid butter; rancid bacon; sour milk*
A1PerfDegrad(milk) = *sour*
A1PerfDegrad(butter) = *rancid*

There is some kind of simile relationship existing between adjective and noun: such as in *cherubic smile; beady eye; watery moon*
Figur(smile) = *cherubic*
Bon(smile) = *cherubic*
Figur(eye) = *beady*

The prototypical properties of a noun are often expressed by the adjective: such as in *intrepid explorer; ancient custom*
Qual1(offend) = *violent*
Magn(ovation) = *standing*
ver(instrument) = *precise*

The prototypical attitudes are commonly expressed through the modifying adjective: such as in *unfortunate accident; idle gossip*
Magn(observer) = *close*
Bon(observer) = *keen*

Some subscripts and superscripts have been added in order to provide a finer-grained description on paradigmatic and syntagmatic relations (see (Wanner, 1996)). Below are some examples:

Subscripts
semantic components:
$\textbf{Magn}_{csq}$(illness) = serious

Superscripts
temporal/quantitative:
$\textbf{Magn}^{temp}$(experience) = lengthy
$\textbf{Magn}^{quant}$(experience) = considerable

The MTT approach is very interesting as it provides a model of production well suited for generation with its different strata and also a lot of lexical-semantic information. All the collocational information is listed in a static way. One of the main drawbacks of the approach is the lack of any predictable calculi on the possible expressions which can collocate with each other **semantically**.

---

[3]There are about 60 LFs listed said to be universal; the lexicographic approach of Mel'čuk and Žolkovsky has been applied among other languages to Russian, French, German and English.

3.0.6.   The Conceptual Approach

The approach described in (Viegas, 1998) builds on (Nirenburg *et al.*, 1992) from the semantic viewpoint as it uses an ontology of concepts to describe the meanings of the words. More specifically, it uses the MikroKosmos ontology (Mahesh, 1996).

In order to account for the continuum we find in natural languages, one needs to account for a continuum perspective, spanning the range from free-combining words to idioms, with semantic collocations and idiosyncrasies in between as defined in (Viegas, 1998):

- **free-combining words** (*the girl ate candies*)

- **semantic collocations** (*fast car; long book*)[4]

- **restricted semantic co-occurrence**, where the meaning of the co-occurrence is semi-compositional between the base and the collocate (*strong coffee, pay attention, heavy smoker, ...*)

- **restricted lexical co-occurrence**, where the meaning of the collocate is compositional but has a lexical idiosyncratic behaviour (*lecture ... student; rancid butter; sour milk*).

- **idioms** (*to kick the (proverbial) bucket*)

Formally, one goes from a purely compositional approach in "free-combining words" to a non-compositional approach in idioms. In between, a (semi-)compositional approach is still possible. (Viegas and Bouillon, 1994) showed that the set of what are conventionally considered as idiosyncrasies can be reduced by differentiating "true" idiosyncrasies (difficult to derive or calculate) from expressions which have well-defined calculi, being compositional in nature, and that have been called semantic collocations.

Semantic collocations (or co-occurrences) refer to co-occurrences which are compositional in nature. For instance, in *fast car*, *fast* is an event modifier. In this case, it modifies one prototypical property of *car*, which can be coded directly in the lexicon (as *run* in the case of (Pustejovky, 1995), in the telic role of the qualia), or in 14 the world knowledge (as "motion-event" as is the case in (Viegas, 1998)). In any case, it is possible to analyse and generate the semantic co-occurrence by compositional means only. Here, one only needs to specify which words can co-occur **semantically**, with no needs to restrict which words can co-occur lexically (in that sense, the conceptual approach might be more general than the generative lexicon approach); whereas in the case of restricted co-occurrences, as defined below, one needs to specify which words can co-occur **lexically**, whether the co-occurrence is treated as being semi-compositional or fully compositional.

Restricted semantic co-occurrence   The semantics of the combination of the entries is semi-compositional. In other words, there is an entry in the lexicon for the base, (the semantic collocate is encoded inside the base), whereas one cannot directly refer to the sense of the semantic collocate in the lexicon, as it is not part of its senses. The co-occurrence is assigned a new semi-compositional sense, where the sense of the base is composed with a new sense for the collocate.

---

[4]See (Viegas and Bouillon, 1994) for an account of such expressions using a coercion operator.

```
#0=[key: "smoker",
    co-occur: [syntagmatic: LSFIntensity
               [base: #0, collocate:
                      [key: "heavy",
                       gram: [subCat: Attributive,
                       freq: [value: 8]]]]] ...]
```

Here, the Lexico-Semantic-Function (LSF)s (e.g., LSFIntensity) are equivalent (and some identical) to the LFs provided in the ECD. The notion of LSF is the same as that of LFs. However, LSFs and LFs are different in two ways: i) conceptually, LSFs are organised into an inheritance hierarchy; ii) formally, they are rules, and produce a new entry composed of two entries, the base with the collocate. As such, the new composed entry is ready for processing. These LSFs signal a compositional syntax and a semi-compositional semantics. For instance, a *heavy smoker* is somebody who smokes a lot, and not a *fat* person. It has been shown that one cannot code in the lexicon all uses of *heavy* for *heavy smoker, heavy drinker, ...* Therefore, there is not in the lexicon entry for *heavy* a sense for *a lot*, or a sense for *strong* to be composed with *wine*, etc. It is well known that such co-occurrences are lexically marked; if one allowed in the lexicons a proliferation of senses, multiplying ambiguities in analysis and choices in generation, then there would be no limit to what could be combined and the system could end up generating **heavy coffee* with the sense of *strong* for *heavy*, in the lexicon.

The left hand-side of the rule LSFIntensity specifies an "Intensity-Attribute" applied to an event which accepts aspectual features of duration. The event is here *smoke*. The LSFIntensity also provides the syntax-semantic interface, allowing for an Adj-Noun construction to be either predicative (*the car is red*) or attributive (*the red car*). The co-occurrence needs therefore to be restricted to the Attributive use only, as the predicative use is not allowed: (*the smoker is heavy*) has a literal meaning or figurative, but not collocational.

Restricted lexical co-occurrence   The semantics of the combination of the entries is compositional. In other words, there are entries in the lexicon for the base and the collocate, with the same senses as in the co-occurrence. Therefore, one can directly refer to the senses of the co-occurring words. What is being captured here is a lexical idiosyncrasy or in other words, one specifies that the system should prefer this particular combination of words. This is useful for analysis, where it can help disambiguate a sense, and is most relevant for generation; it can be viewed as a preference among the paradigmatic family of the co-occurrence.

```
#0=[key: "truth",
    co-occur: [syntagmatic: LSFSyn
               [base: #0, collocate:
                      [key: "plain",
                       sense: adj2,
                       lex-rul: [comp:no,
                                 superl:no]]]] ...]
```

The LSFSyn produces here a new entry composed of two or more entries. As such, the new entry is ready for processing. LSFSyn signals a compositional syntax and a compositional semantics, and restricts the use of lexemes to be used in the composition. One can directly refer to the sense of the collocate, as it is part of the lexicon.

The entry for *truth* specifies one co-occurrence (*plain truth*), where the sense of *plain* here is adj2 (obvious), and not say adj3 (flat). The syntagmatic expression inherits all the zones of the entry for *plain*, sense adj2, one only needs to code here the irregularities. For instance, *plain* can be used as *plainer, plainest* in its *plain* sense in its adj2 entry, but not as such within the lexical co-occurrence **plainer truth*, **plainest truth*, it must therefore be blocked in the collocate, as expressed in (comp: no, superl: no). In other words, *plainer/plainest truth* will not be generated.

Multinguality in Syntagmatic Relations   The above approach has been generalized to be used in a multilingual environment. Fo instance, in the case of Quantifiable Events, one can expect, for a same family of languages, that co-occurrences behave in the same way. Let us consider the examples below:

French: *grand fumeur; gros mangeur*
Spanish: *gran fumador; gran comelon*
English: *heavy smoker; big eater*

Assuming there is a database of LSFs (e.g., LSFIntensity which applies to quantifiable events), then SMOKE will inherit LSFIntensity. Assuming now that the Spanish collocate *gran* has been manually acquired for "fumador" (smoker), the system could guess and attach the French translation "grand" of "gran" to the base "fumeur" (smoker).

Note that manual checking is required as for instance *eater* has different collocates in French and Spanish, "grand/gros" and "gran", respectively.

4. Papers Workshop Presentation

The articles included in this proceedings deal with various aspects of adjective polymorphism, from the following perspective: lexicographic, syntactic, logical, semantic and morphological. Some articles provide descriptive studies only (e.g., Fabienne Martin and Beatrice Warren). Other articles provide representations, encoding adjective properties, using formalisms well known in NLP: syntactic (HPSG) (e.g. Anne Abeillé and Danièle Godard); semantic (Generative Lexicon) (e.g. Toni Badia and Roser Saurì). From a syntactic viewpoint, one of the major problems for romance languages is to explain syntactic differences, such as prenominal or postnominal positions (e.g. Mario Saltarelli, Christopher Laenzlinger and Thierry Etchegoyhen). From a semantic viewpoint, the focus of the research is on deriving the meanings and/or the syntax of adjectives based on the lexical semantic properties of the noun and the adjective (e.g. Pablo Gamallo and Marie-Laure Reinberger, Kyoko Kanzaki and Hitoshi Isahara). How do adjective relate to verbs or nouns is also an issue which can help in identifying the semantics of the adjective (e.g. Christopher Kennedy and Louise McNally) and also in acquiring them (e.g. Kadri Muischnek, Kaili Müürisep and Tiina Puolakainen).

The Mysterious Bona Fide Adjectives - *Beatrice Warren*

> This paper addresses the distinction between bona fide adjectives and non-predicating adjectives. Bona fide adjectives can occur in attributive as well as predicative positions: *the dirty rag/the rag is dirty*; they can be graded: *very dirty/dirtier/ dirtiest*; they can be paratactically related to other adjectives: *dirty, nasty rag/ nasty, dirty rag*. Non-predicating adjectives, however, are constrained to attributive position: *musical instrument/*the instrument is musical*; they are non-gradable: **very musical instrument*, and they do not tolerate separation from their heads: **musical expensive instrument*. In this paper it is argued that the main difference between predicating and non-predicating adjectives depends on the function of the adjective. Moreover, a distinction, along morphological and semantic lines, is made between bona fide adjectives and non-predicating ones.

"Elle est agaçante d'entrer sans frapper": les adjectifs verbaux émotifs présents avec infinitive en français - *Fabienne Martin*

> In this paper, the author studies infinitive French clauses which include deverbal adjectives, ending in -ANT (-ING), in structures such as [ADJECTIF de V]. The author focuses on the emotional status of these adjectives. ¿From a corpus of 300 verbal emotive present adjectives ('AV-ant'), and shows how some intentional, generic, and evidential factors can affect the possibility, for these adjectives, to receive a Vcomp. The author distinguished two uses of the infinitive, one instantiating the obvious cause of emotion of the experiencer, the other categorizing an evaluative inclination of the AV-ant subject. Are also considered in this paper certain coreference problems between thematic roles of the 'AV-ant' and the infinitive.

La place de l'adjectif épithète en français: modélisation en HPSG - *Anne Abeillé and Danièle Godard*

> This paper proposes to formalize Adjective-Noun constructions for French, with a focus on syntactic constraints. It is important in French to know the position of the adjective

(postnominal or prenominal) as they can exhibit different meanings. The authors model the data using HPSG making the following assumptions: -there is no one-to-one mapping between structure and word order; -when multiple positions are possible, all are made available, in parallel, in the grammar, without any movement rule from either an initial or canonical position. Only competence factors has been taken into account in this paper, meaning that all grammatical orders are left equal. Room is left for adding systems of preferences.

An Integrated Theory of Adjectives - *Mario Saltarelli*

This paper addresses adjectives in Romance languages. The Romance tradition recognizes an attributive sense for adjectives used in prenominal position in a noun phrase: in Italian, "un pover'uomo/un uomo povero" or in Romanian "sa(racul om/omul sa(rac" (a poor (pitiable) man/ a poor (indigent) man). In this paper the author proposes an integrated theory of the semantics and syntax of attributive and predicative adjectives. Here adjective phrases (AP)s are treated as projections of A0 which merge with NP either as complements of D' or of N'. Predicate adjectives are assumed to be "unsaturated." Therefore, the lexical properties of adjectives which merge as complements of N' are referentially saturated by identification with the head noun, hence the predicative interpretation. In contrast, adjectives which are complements of D' are not referentially saturated, hence the variable, speaker oriented, quasi-metaphorical reading of the lexical properties of the attributive adjective.

Le traitement des adjectifs dans le groupe nominal - *Christopher Laenzlinger and Thierry Etchegoyhen*

This paper addresses French attributive adjectives. As seen in previous papers, the order of the adjective (prenominal or postnominal) can have various interpretations. Moreover there are also sequential constraints between adjectives. Adjectives are studied along typological distinctions. The author presents lexico-semantic features helping in determining the position of the adjective. This study aims at identifying features and rules helping in analyzing and generating French attributive adjective-noun constructions.

Semantic Disambiguation of Adjectives in Local Context: a Generative Approach - *Toni Badia and Roser Saurì*

In this paper the authors concentrate on the semantic disambiguation of adjectives using strictly local context information, i.e., by dealing with the semantic relation between adjectives and their nominal head. They focus on some cases that have been standardly dealt with as collocations by using a list of separate entries. They also show that the Generative Lexicon approach (Pustejovsky, 1995) is adequate, provided that some fine-grained distinctions are introduced, such as a classification of nouns, a classification of adjectives and a developed set of selectional restrictions.

Activation de l'information lexicale dans la combinaison adjectif-nom - *Pablo Gamallo and Marie-Laure Reinberger*

This paper presents elements to help organize lexical structures which carry some degree of indetermination. The authors use compositionality as a means to regulate the access to the information found in lexical structures. Compositionality is a the basis of the dynamically building the meaning. This work simulates the interpretation of "adjective-noun" constructions.

Several aspects of semantic relations between adjectives and their head nouns in Japanese - *Kyoko Kanzaki and Hitoshi Isahara*

This paper addresses adjectives in Japanese. The authors study paraphrastic relationships between attributive and predicative positions to help them build a lexical classification. The relation between the adjective and the head noun include feature-value type information which relates the adjective to the head noun and also contextual-type information. The authors discuss various semantic relations between adjectives and nouns in Japanese.

Degree Modification and the Scalar Structure of Gradable Adjectives - *Christopher Kennedy and Louise McNally*

This papers addresses the interaction of adjectives and degree modifiers such as *very* and *well*. Focusing on the case of deverbal gradable adjectives, the authors demonstrate that the distribution of degree modifiers is closely tied to the scalar structure of the adjectives they modify, and that scale structure also determines one of the core semantic properties of gradable adjectives: the nature of the "standard values" according to which sentences involving adjectival predications are judged to be true. In the case of deverbal gradable adjectives, scalar structure can be inferred from the aspectual properties of the source verbs.

Automatic Analysis of Adjectives in Estonian - *Kadri Muischnek, Kaili Müürisep and Tiina Puolakainen*

In this paper, the authors describe problems that were encountered in the morphological and syntactic analysis of adjectives in Estonian and give an overview of the semantic classes of adjectives which should be distinguished in order to arrive at correct automatic analysis. On the basis of their present experience, the authors argue that for a successful analysis of adjectives in Estonian, no additional thorough lexical database is needed. It is sufficient to examine the context, to consider suffixes and to draw up a few lists.

References

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

Bouillon, P., Polymorphie et sémantique lexicale, Lille: Presse Universitaire du Septentrion, 1997.

Bouillon, P., The adjective *vieux*: The point of view of *Generative Lexicon*, in: E. Viegas (ed), *Breadth and Depth of Semantic Lexicons*, Kluwer, in press.

Busa, F., *Compositionality and the semantics of nominals*, PhD dissertation, Brandeis University, 1996.

Carlson, G.N., A Unified Analysis of the English Bare Plural, in : *Linguistics and Philosophy*, 1977, pp. 413-455.

Carlson, G.N. and F. J. Pelletier, *The Generic Book*, Chicago and Londres: Chicago University Press, 1995.

Chierchia, G. and S. McConnell-Ginet, *An Introduction to Semantics*, Cambridge:MIT, 1990.

Dixon, R. M. W., *A new approach to English Grammar, on Semantic Principles*, Oxford:Clarendon Press, 1991.

Gross, D. and K.J. Miller, Adjectives in *WordNet*, in :*Lexic ography*, 3:4, 1990, pp. 265-278.

Iordanskaja, L., R., Kittredge and A. Polguère. Lexical Selection and Paraphrase in a Meaning-text Generation Model. In C. L. Paris, W. Swartout and W. Mann (eds), *NLG in AI and CL*. Kluwer Academic Publishers, 1991.

Lahav, R., Against compositionality: the case of adjectives, in: *Philosophical studies*, 57, 1989.

Levi, J.N., *The syntax and semantics of complex nominals*.- New York:Academic Press, 1978.

Mahesh, K. *Ontology Development: Ideology and Methodology*. Technical Report MCCS-96-292, Computing Research Laboratory, New Mexico State University, 1996.

Marx, W., The meaning-confining function of the adjective, in: Rickheit et Bock (éds), 1983.

Mel'čuk, I. Paraphrase et lexique dans la théorie Sens-Texte. In Bes and Fuchs (ed.) *Lexique 6*, 1988.

Nirenburg, S. and I. Nirenburg. A Framework for Lexical Selection in NLG. In *COLING 88*, 1988.

Nirenburg, S., J. Carbonell, M. Tomita, M. and K. Goodman. *Machine Translation: A Knowledge-based Approach*. Selection in NLG. Morgan Kaufmann Publishers, 1992.

Parsons, T., *Event in the semantics of english: A study in subatom ic semantics*, Cambridge:MIT Press, 1990.

Picabia, L., *Les constructions adjectivales en français. Systématique transformationnelle*, Genève, Paris : Droz, 1978.

Pustejovky, J., *The Generative Lexicon*, Cambridge:MIT Press, 1995

Quirk, R., S. Greenbaum, G. Leech and J. Svartvik, *A Comprehensiv e Grammar of the English Language*, London et New York:Longman, 1994, (Twelfth impression).

Ramos, M., A. Tutin and G. Lapalme. Lexical Functions of Explanatory Combinatorial Dictionary for Lexicalization in Text Generation. In P. St-Dizier and E. Viegas (Ed) *Computational Lexical Semantics*. CUP, 1995.

Raskin, V. and S. Nirenburg,*Lexical Semantics of Adjectives. A microtheory of Adjectival meaning*, in: MCCS-95-288, 1995.

Raskin, V. and S. Nirenburg, Adjectival modification in text meaning representation. In the Proceedings of *Coling-96*, Copenhagen, pp. 842-847, 1996.

Rickheit, G. and M. Bock, *Psycholinguistic Studies in Language Processing*, Walter de Gruyter:Berlin, 1983.

Sadler, L. and D. Arnold, *Prenominal Adjectives and the Phrasal/ Lexical Distinction*, mss., 1993.

Silva, G. and S.A. Thompson, "On the syntax of adjectives with 'it' subject and infinitival complements in English", in : *Studies in Langua ge*, 1:1, 1977, pp. 109-126.

Viegas, E. and P. Bouillon. Semantic Lexicons: the Cornerstone for Lexical Choice in Natural Language Generation. In Proceedings of the *7th International Workshop in Natural Language Generation*, Kennebunkport, 1994.

Viegas, E. Multilinguality and Genericity in Syntagmatic Relations. In Proc. of *the Second International Workshop on Multilinguality in the Lexicon*, as part of *the 13th European Conference in Artificial Intelligence*, Brighton, UK, 1998.

Vendler, Z., *Adjectives and Nominalization*, La Haye: Mouton, 1968.

Wanner, L. *Lexical Functions in Lexicography and Natural Language Processing*. John Benjamin Publishing Company, 1996.