Pierrette Bouillon
ISSCO, University of Geneva

Evelyne Viegas
Computer Science Department, Brandeis University, Waltham

A Semi–Polymorphic Approach to the Interpretation of Adjectival Constructions: A Cross–Linguistic Perspective

Abstract

In this paper, we investigate how the field of machine translation (MT) can benefit from a semi–polymorphic approach to characterize natural languages. The approach taken here is in contrast to the kind of word–sense enumeration found in dictionaries: we show that it is possible to minimize the problems entailed by mismatches in MT if we are first able to capture ambiguity among polysemous lexical items. Our goal is twofold: from a theoretical viewpoint, we can account for the creative power of word senses and explain the semantic relatedness of polysemous lexical items; from a practical point of view in MT, we can detect and treat ambiguity in the source language and then translate to the target language in a way which deals with this ambiguity. We illustrate our point with reference to the interpretation of adjectival constructions.

1. Problems for MT: polysemy and mismatches

Translating one language into another one is a difficult task partly because of lexical mismatches; a lexical item in the source language (SL) can have more than one translation in the target language (TL) and vice versa (Hutchins et al. 1992). Our main goal is to show that many of the problems accounted for in terms of mismatches can be solved if we start reducing ambiguity among "polysemous" lexical items in **each language**. For instance, how does one translate *an old book* in Spanish where we have two lexical items *viejo* referring to the physical aspect of *book* and *antiguo* referring to the text, a book written long ago? It is possible to list all the "different senses" in different entries (e.g. *book1, book2*), as would be the case in a lexicographic approach. However, apart from increasing the size of the lexicon in a rather random way, a mere lexicographic approach fails to account for the semantic relatedness of lexical items of the type of *book*. Furthermore, if we take this view for polysemous adjectives when in composition with polysemous nouns, arriving at the correct interpretation requires stipulating at best one rule per combination. Here, we go beyond a descriptive approach, adopting a computational linguistics perspective, in

which we can explain the behaviour of these lexical items within sentences in a given language, by enriching the semantics of lexical items. In 1.1 and 1.2 we illustrate cases of polysemy in nouns and adjective–noun constructions respectively, and show why previous approaches fail to give a homogeneous account of such cases. In 2. we present our approach which allows us to give a proper unified treatment of polysemous nominals in composition with adjectives. In 3. we illustrate through a case study how the field of machine translation may benefit from our approach.

1.1 The polysemy of nominals

For the purpose of this paper we will concentrate on "covert" relational nominals[1] belonging to the class of artifacts, such as *book, novel, record, door* that exhibit semantic polysemy. We argue that these cases have actually well–defined calculi. If we look at examples (1a) through (1c):

(1a) *This book is heavy to carry around.* (physical object)
(1b) *I read an angry book.* (text)
(1c) *This book is great!* (text and/or physical object)

(1a) and (1b) illustrate the polysemy between, on one hand, the physical object and on the other hand, the notion of text, whereas (1c) can either refer to one of the aspects or to both aspects within the same sentence.

Traditional approaches, from transformational grammars to classical Montague grammars, account for this lexical ambiguity by postulating different entries per lexical item; this fails to capture the core semantics of the lexical items, leaving the *complementary*[2] senses unrelated. In 2.2 we present our analysis of nominals which is in keeping with the work of Pustejovsky (1994b): we suggest that covert relational nominals should have a relational structure, thus capturing polysemy within the lexical structure.

1.2 Adjectival polysemy

Adjectival polysemy (examples (2)), moreso than nominals, is so clear that adjectives have been used to argue for the infinite variance of language (Rastier 1987: 71):

(2a) *un vieux livre*   –> worn (absolute)
(2b) *un vieil ami*   –> aged (relative)
(2c)   friend of long standing (property modifying)

First, it is well known that the same adjective can belong to different semantic classes – absolute (abs.), relative (rel.) and property modifying (pm., in the terms of Arnold (1989) – which differ in their logical behaviour. (2a) is intersective and implies that the entity is *a book* and is *worn*, while (2b)

and (2c) are not: (2b) supposes that the entity is *a man*, who is *old for a man* while (2c) results from an interaction of the senses of the nouns and the adjective: the man need not be old (Arnold 1989 and Siegel 1980). Secondly, it has often been noticed that relative and property modifying adjectives have different interpretations depending on the nouns.

Two main questions have been addressed in the literature: how to deal with this logical behaviour, and how to predict the interpretations of relative and property modifying adjectives? As in the case of nouns, treatments based on transformational grammar or Montague analyses, when postulating different transformational origins or two different categories of adjectives, are of little interest: first, they are based on a special, but questionable link between intersective sense/predicative position and non–intersective sense/attributive position and second, they cannot relate the semantics of the two kinds of adjectives (Arnold 1980).

In this work, we take a different approach, following more the works on adjectives in cognitive grammar (Taylor 1992), in interpretative semantics (Rastier 1987) and in psycholinguistics (Miller and Johnson–Laird 1976): we postulate that the ambiguity of adjectives results not from its polysemy or from the syntax, but from the structural semantic representation of the noun being modified by the adjective.

2. Capturing ambiguity: a semi–polymorphic approach

2.1 A well–suited approach to treat polysemy: the generative lexicon

We discussed in the first section how a sense enumeration approach to polysemy fails to capture the semantic relatedness of *complementary* senses. The multiple listing of a word is well expressed by monomorphic languages where each lexical item is assigned a single type and denotation. The opposite view is held by unrestricted polymorphic languages where the lexical item has no literal meaning *per se*: the meaning of the words in context is constrained by some "background knowledge". The approach taken here is that suggested by (Pustejovsky 1994a), within the framework of his "Generative Lexicon" (GL), where he introduces "semi–polymorphic" languages. The idea is that all lexical items are semantically active and have a richly typed semantic representation.

Briefly, the GL is a system which involves four levels of representation, namely the **argument structure** which specifies the predicate argument structure for a word and the conditions under which the variables map to syntactic expressions; the **event structure** which gives the particular event types such as S (state), P (process) or T (transition); the **qualia structure** and **inheritance structure** developed below:

Qualia structure.
The qualia structure gives the essential characteristics of the lexical item under consideration distributed among four roles,

- Formal: allows the object to be distinguished within a larger domain;
- Constitutive: specifies the relation between the object and its constituents or parts;
- Telic: gives the purpose and function of the object;
- Agentive: specifies what brings the object about.

Inheritance structure
The inheritance structure involves two different kinds of mechanisms:

- the *fixed* inheritance mechanism which is basically a fixed network of the traditional *isa* relationship found in AI, enriched with the different roles of the qualia structure;
- the *projective* inheritance mechanism which can be intuitively characterized as a way of triggering semantically related concepts which define for each role the projective conclusion space (*PCS*). For instance, in the *PCS* of the telic and agentive roles of *book*, we will find at least the following predicates: *read, reissue, annotate,...* and *write, print, bind, ...* (respectively).

Another important notion introduced is the notion of *lexical conceptual paradigms (LCPs)*, as formalized in (Pustejovsky 1994b). We will say that the aim of an LCP is to capture the conceptual regularities across languages in terms of cognitive invariants, like "physical–object", "aperture", "natural kind" and alternations such as "container/containee", etc. For the study of the invariants we will follow Dubois and Pereita (1993) in their analysis of categorisation in relation with cognition. Moreover, depending on the combination of LCPs the lexical item enters into, it will exhibit different linguistic behaviours at the level of natural languages. For instance, we can insert a leaflet *in* a book (as book is an *information–physical_object–container–LCP*), but we can neither insert nor put a carpet *in* the floor. We exemplify these notions below.

2.2 Modeling constraints for adjectival interpretation

2.2.1 The representation of nominals

The lexical entry of covert relational nominals belonging to the class of artifacts must contain a relational structure, as mentioned in Section 1.1, in

order to be able to account for the different examples listed in 1.1 (examples 1). Moreover, it enables us to derive in a homogeneous way the right

interpretation of polysemous adjectives in composition with that kind of nominals as will be seen in the next paragraph. For the purpose of this paper and in order to exemplify the notions presented in 2.1, we will only give a partial representation of *book* below[3]:

$$
\begin{bmatrix}
\text{book} \\[4pt]
\text{ARGSTR} = \begin{bmatrix} \text{ARG1} = x:\text{text} \\ \text{ARG2} = y:\text{paper} \end{bmatrix} \\[12pt]
\text{QUALIA} = \begin{bmatrix} \text{information-phys\_obj-container-LCP} \\ \text{FORM} = \text{book\_hold}(y,x) \\ \text{TELIC} = \text{read}(T,v:\text{individual},x), \ \text{publish}(T,v:\text{publisher},y) \\ \text{AGENT} = \text{write}(T,u:\text{writer},x), \ \text{print}(T,z:\text{printer},y) \end{bmatrix}
\end{bmatrix}
$$

Briefly, this states that book inherits from the qualia of the relational text–physical_object–container–*LCP*, although imposing additional constraints on its own, represented here as the two arguments[4], namely ARG1 and ARG2: *text* and *paper* respectively. Moreover, we have specified two defaults for the telic and agentive roles, each refering to one aspect of *book*, either **text** or **paper**. The sorts *publisher, writer, printer* are organized hierarchically with *individual* as a common super–type. This nominal representation, over which operate the projective operators, thus dynamically generating the predicates of the PCS (see 2.1), enables us to capture all the *complementary* nominal polysemies: *The writer began his third book* (writing), my sister began "The Bostonians" (reading); the binder finished the books for CUP (binding), etc. We now turn in the next section to the adjectival interpretation of polysemous adjectives in composition with nouns.

2.2.2 Adjectival interpretation

Within the semi–polymorphic approach taken here, the adjectives will submodify the different qualia roles and the arguments inside them and, depending on the information they modify, will acquire their different senses. In general terms, an abs. adjective will always modify the formal role of a noun (example (3)), a pm. the agentive or the telic (examples (5ab)) while a rel. seems able to modify the formal (5c) or the agentive (9b), a fact that can explain the controversial status of this semantic class, often conflated with the two others (Arnold 1989: 108):

    (3) *un livre rouge*   –> the physical_obj is red (modification of the
                                          formal role)

> (4a) *un livre triste* -> to read which makes one sad (modification of the individual inside the telic role)
>
> (4b) -> whose writing expresses the sadness of the writer (modification of the individual inside the agentive role)
>
> (4c) *un livre anxieux* -> whose writing expresses the anxiety of the writer (modification of the individual inside the agentive role)
>
> (4d) *une robe triste* -> *that causes the sadness of the person who wears it
>
> (5a) *un long livre* -> taking a long time to read (modification of the telic role)
>
> (5b) -> taking a long time to write (modification of the agentive role)
>
> (5c) -> which has a long size (modification of the formal role)

The interest of this approach is that the richness of typing can explain why some adjectives seem to license the modification of a non–selected type. *Triste*, for example, selects for a noun of type *individual*, but is possible with *livre* in (4a), because the type required by the adjective *triste* (i.e. *individual*) is present in the qualia of the noun, which makes explicit different relations between the type selected by the adjective *person* and the type of the noun (text or physical object), in this case how a person can use or produce the object *livre*.

This analysis raises two important issues that we cannot develop here and need further elaboration:

- The definition and formalization of morphological syntactic, semantic, pragmatic and cognitive constraints which explain why, within a particular class, some adjectives allow modification of different roles and others do not. In (4abc), for example, cognitive principles can explain why *triste* can modify the agentive and telic roles, while *anxieux* can modify only the agentive: while *triste* is a communicable feeling, this is not the case with anxiety. Similarly, (4a) is possible, but not (4d) because there must be a direct causal link between the event expressed in the agentive/telic role and the sadness of the individual. This link does not relate in our societies sadness and the wearing of a particular dress.

- The choice of a particular sense in context. The selection of a particular quale by an adjective can be influenced by the other words of the sentence, the position of adjectives as in French and Spanish, possessive determiners and typography (hyphen or quotes, which indicate an interaction of the two senses). In sentence (6), for example,

the word *mourir* (die) implies rel. sense *aged* (modification of the agentive) and in sentence (7), the hyphen between the noun and the adjective implies pm. sense (in this case, modification of the telic role – *of long standing*):

(6) ... *des vieux musiciens ... disposés à transmettre leur savoir avant de mourir.*

(7) ... *et l'intégrisme des vieux–croyants ...*

We will see now how the kind of interpretation of adjective–noun sequences given in (3), (4) and (5) can be used to obtain better translations.

3. A case study for MT: the translation of *vieux*

It has often been said that a complete disambiguation/understanding of a SL is not necessary for adequate translation in TL (for a dissenting view, see Nirenburg (1993)). The translation of French and Spanish adjectives show the contrary, as we will see for *vieux*:

(8a) *un vieil ami*   (aged)   –> old friend –> *un amigo viejo*
(8b) (friend of long standing)   –> old friend –> *un viejo amigo*
(8c) *un vieux musicien* (aged)   –> old musician –> *un musico viejo*
(8d) (musician of long standing)   –> musician of long standing –> *un viejo musico*

(9a) *un vieux livre* (old)   –> old book –> *un libro viejo*
(9b) (written long ago)   –> old book –> *un libro antiguo*

As these examples show, it is possible to translate the French adjective *vieux* (rel. or pm.) into English as, among other things, *old* (rel. or pm.) and *of long standing* (pm.), the correct translation depending on the different lexical interpretations of the adjective–noun pairs in the SL and TL languages. Two cases of translation can be illustrated.

Adjective–noun pairs in the SL and TL give rise to the same ambiguity in terms of GL. *Un vieil ami* and *an old friend*, for example, receive the same GL interpretations, one where *vieux* modifies the agentive quale (born a long time ago, aged) and the other where it modifies the telic quale (modification of the relational state of friend). As there is a one to one mapping in the two languages, the translation is straightforward.

The ambiguity is not shared by the different languages: in order to recognize the mismatch, it is necessary to know the potential readings of the source and target adjective–noun pairs. For example, *un vieux musicien* and *an old musician* do not share the same interpretations: the former has the

same GL interpretations as *un vieil ami* — modification of the agentive role (born a long time ago, aged) and the telic (who has been playing music for a long time) — while the latter can only receive the relative sense (aged). This is because the English adjective *old* can only get the sense *of long standing, though not necessarily old* with a relational state in the telic. As a result of this mismatch, *old* cannot be used to express the two senses and *un vieux musicien* should therefore receive two translations, one with *old: an old musician* and a second with e.g. *of long standing: a musician of long standing*. The GL representations are also very useful when we translate between Spanish and French: in order to know that un *vieil ami* must receive two translations in Spanish (with the adjective before and after the noun), we must know that the French adjective is ambiguous and is able to modify the telic and the agentive role.

This suggests an interesting perspective for the future. Thanks to a better representation of the logical polysemy of noun phrases, it would be possible to envisage a more ambitious approach to translation, which would try to save, when possible, the logical polysemy in the target language. In other words, if one NP has two logical readings, as *un vieil ami* in French, it must be translated, if possible, as a target noun phrase which shares the same meanings. If this target word does not exist, then it is necessary to generate two translations or to use knowledge about context/situation to dis– ambiguate the source noun phrase.

4. Conclusion

In this article, we emphasized the interest of a semi–polymorphic approach to the treatment of polysemy. In the context of MT and, more precisely, of the translation of complex NPs, this kind of approach has many advantages: on the theoretical side, it enables us to account for the creative power of word senses and explain the semantic relatedness of logical polysemous items; on a practical side, first, it enables us to reduce the size of the lexicon, second, it enables us to find the ambiguity of adj–noun sequences in the source language and translate them in a way which deals with this ambiguity.

Notes

1   For a broader account of the semantic interpretation of polysemous nominals, including nominalizations, see Pustejovsky and Anick (1988). Besides, we use "covert" to differentiate traditional relational nominals (such as *friend, father, cousin*), from the class of nouns which exhibit a polysemous behaviour (such as *book, door, record*).
2   Weinreich (1964) makes the distinction between *contrastive* and *complementary* ambiguity: the former deals with homonymous nouns such as *record*, a written statement of facts or events and the gramophone record or disc, whereas the latter deals with complementary aspects such as the physical object and the music for *record*.
3   We mainly use the approach to typed feature structures as described in Carpenter (1992). Besides, we cannot develop in this paper the way the information is inherited in this partial

lexical entry. Briefly, we can build the class of concepts *book* belongs to and also the specific sortal/type constraints brought to this instance.
4   The arguments, ARG1 and ARG2, are to be taken as logical parameters providing type information for lexical items as discussed in Pustejovsky (1994a).

References

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

Carpenter, B. 1992. *The Logic of Typed Feature Structures*. Cambridge: CUP.

Dubois, D., Peraita, H. 1993. "Ontology for Concepts and Categories: Is there any difference? A few Empirical Data and Questions for an Ontology of Categorial Knowledge". In Proceedings *Formal Ontology*, Padova, March 17–19.

Hutchins, W. J., Somers, H.L. 1992. *An introduction to Machine Translation*, London: Academic Press.

Miller, G. A., Johnson–Laird, P. N. 1976. *Language and Perception*. Cambridge, MA: Belknap.

Nirenburg, S. 1993. "L'interlangue et le traitement du sens en TA". In P. Bouillon et A. Clas (Eds). *La traductique*. Montréal: PUM.

Pustejovsky, J., Anick, P. 1988. "On the Semantic Interpretation of Nominals". In *Coling 1988*, vol.2: 518–523.

Pustejovsky, J. 1994a. "Linguistic Constraints on Type Coercion". In P. Saint–Dizier, E. Viegas (Eds). *Computational Lexical Semantics*. Cambridge: Cambridge University Press.

Pustejovsky, J. 1994b. "Semantic Typing and Degrees of Polymorphism". In C. Martin–Vide (Ed). *Current Issues in Mathematical Linguistics*. Elsevier North Holland Inc.

Rastier, F. 1987. *Sémantique Interprétative*. Paris : Puf.

Siegel, M. 1980. *Capturing the adjective*. New York: Garland Publishing.

Taylor, J. R. 1992. "Old problems: Adjectives in Cognitive Grammar". In *Cognitive Linguistics*, 3–1: 1–37.

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

Weinreich. 1964. "Webster's Third: A Critique of its Semantics". In *International Journal of American Linguistics* 30: 405–409.