

# **The Predicate-Argument Structure**

T.Badia, C.Colominas

Sèrie Monografies, 4

Barcelona  
Universitat Pompeu Fabra. Institut Universitari de Lingüística Aplicada  
1997

Direcció de les Publicacions de l'IULA: M. Teresa Cabré

Primera edició: 1997

© els autors

© Institut Universitari de Lingüística Aplicada

La Rambla, 30-32

08002 Barcelona

Dipòsit legal: B-34.280-2002

# The Predicate-Argument Structure

Toni Badia, Carme Colominas  
Institut Universitari de Lingüística Aplicada  
Universitat Pompeu Fabra  
tbadia@upf.es , colomina@upf.es  
Project Number: MLAP 93-15

## **Abstract**

In this paper we present a proposal for the representation of the Predicate-Argument Structure in typed feature-structure formalisms. After a general discussion of the aims and the level of description of the proposed representation, the paper focusses on a thorough exemplification of the Predicate-Argument Structures for all major categories (verbs, adjectives, prepositions and nouns) and of some modification relations.

## **Abstract**

Aquest article conté una proposta per a la representació de les estructures predicatives (és a dir, dels predicats amb els seus arguments) en els formalismes basats en estructures de trets tipificades. L'article comença amb una discussió dels objectius i del nivell de descripció de la representació que es proposa; i després se centra en una exemplificació minuciosa de les estructures predicatives de totes les categories majors (verbs, adjectius, preposicions i noms), així com d'algunes relacions de modificació.

# Contents

<b>1</b>	<b>Overview of the Coverage</b>	<b>6</b>
1.1	Introduction . . . . .	6
1.2	The Structure of Predicate and Arguments . . . . .	6
1.3	The Modifiers . . . . .	8
1.4	Formal Issues . . . . .	8
<b>2</b>	<b>Basic Options</b>	<b>10</b>
2.1	Introduction . . . . .	10
2.2	The Basic Structure of the Content of Signs . . . . .	13
2.3	The Labelling of Predicative Types and Arguments . . . . .	17
2.4	Classes of Complements . . . . .	17
2.4.1	Noun Phrases . . . . .	17
2.4.2	Prepositional Phrases . . . . .	18
2.4.3	Other Predicative Complements . . . . .	22
2.4.4	Sentential Complements . . . . .	22
2.5	The Syntactic Representation of Complements . . . . .	22
2.6	The Parallelism between the Syntactic and the Semantic Representation of Complements . . . . .	24
<b>3</b>	<b>Proposal for Formalisation: The Type System</b>	<b>25</b>
3.1	Introduction . . . . .	25
3.2	A Proposal for the PAS Type Hierarchy . . . . .	25
3.3	The Type System Used in the Exemplification . . . . .	27
<b>4</b>	<b>Exemplification: The Lexical Representation of Verbs</b>	<b>28</b>
4.1	Introduction . . . . .	28
4.2	Phrasal Arguments to Verbs . . . . .	29
4.2.1	0-Valent Verbs . . . . .	29
4.2.2	Monovalent Verbs . . . . .	30
4.2.3	Divalent Verbs . . . . .	31
4.2.3.1	Subject-Object — Dynamic . . . . .	31
4.2.3.2	Subject-Object — Stative . . . . .	35

4.2.3.3	Attribute of Subject . . . . .	35
4.2.3.4	Prepositional Verbs . . . . .	37
4.2.3.5	Phrasal Verbs . . . . .	38
4.2.3.6	Arg-Adjunct (Stative) . . . . .	39
4.2.3.7	Arg-Adjunct (Dynamic) . . . . .	40
4.2.3.8	Resultative Secondary Predications . . . . .	42
4.2.3.9	Unaccusative + Arg-adjunct . . . . .	42
4.2.4	Trivalent Verbs . . . . .	43
4.2.4.1	Object + Arg-Adjunct . . . . .	43
4.2.4.2	Object + Arg-Adjunct; relational . . . . .	44
4.2.4.3	Object attributes . . . . .	45
4.2.4.4	Ditransitive Verbs (Two animate participants) . . . . .	46
4.2.4.5	Ditransitive Verbs (Two inanimate participants) . . . . .	47
4.2.4.6	Raising Verbs . . . . .	48
4.2.5	Tetravalent Verbs . . . . .	49
4.3	Clausal Arguments to Verbs . . . . .	49
4.3.1	Divalent Verbs . . . . .	49
4.3.1.1	Finite Complements . . . . .	49
4.3.1.2	Non-finite Complements: Subject Control . . . . .	50
4.3.1.3	Non-finite Complements: Object Raising . . . . .	51
4.3.1.4	Non-finite Complements: Raising to Subject . . . . .	51
4.3.1.5	Non-finite Complements: Small Clause . . . . .	52
4.3.2	Trivalent Verbs . . . . .	52
4.3.2.1	Finite Complements . . . . .	52
4.3.2.2	Non Finite Complements: Subject Control . . . . .	53
4.3.2.3	Non-finite Complements: Indirect Object Control . . . . .	53
4.4	A Note on Modal Verbs . . . . .	54
<b>5</b>	<b>Exemplification: The Adjectival Signs</b>	<b>54</b>
5.1	Introduction . . . . .	54
5.2	Standard, Non-valent Adjectives . . . . .	55
5.3	Valent Adjectives . . . . .	56

<b>6</b>	<b>Exemplification: The Nominal Signs</b>	<b>58</b>
6.1	Introduction . . . . .	58
6.2	Basic Representation for Nominal PAS . . . . .	59
6.3	Predicative Nouns . . . . .	60
6.3.1	Deadjectival Nouns . . . . .	60
6.3.2	Deverbal Nouns . . . . .	61
6.3.2.1	Dynamic Deverbal Nominalisations . . . . .	61
6.3.2.2	Functional Deverbal Nominalisations . . . . .	61
6.4	Simple Nouns . . . . .	63
6.4.1	Abstract Simple Nouns . . . . .	63
6.4.2	Concrete Simple Nouns . . . . .	64
6.4.2.1	Non-argument Constructions . . . . .	64
6.4.2.2	Argument Constructions . . . . .	65
<b>7</b>	<b>Exemplification: The Prepositions</b>	<b>66</b>
<b>8</b>	<b>Exemplification: The Adjuncts</b>	<b>67</b>
8.1	Introduction . . . . .	67
8.2	The Representation of Adjuncts . . . . .	68
8.2.1	The Adjuncts to Nominal Signs . . . . .	69
8.2.2	The Adjuncts to Predicative Signs . . . . .	71
8.3	Non-intersective Adjuncts . . . . .	73
8.4	The Labelling of Adjuncts . . . . .	74

## Introduction

In this paper we present a proposal for the representation of the Predicate-Argument Structure (PAS) in typed feature-structure (TFS) formalisms. The proposal was put forward within the MLAP project number 93-015 “Investigation of linguistic specifications for future industrial standards: The EUROTRA reference manual”.<sup>1</sup> The general purpose of this project was to investigate the suitability of existing linguistic specifications, especially those of the EUROTRA Reference Manual (RM) 7.0 (ET-RM 7, 1992), for their general usefulness in present state-of-the-art NL processing. One of the major outcomes of the project has been the development of a consistent and coherent approach to a wide range of phenomena, in a formalism that does not put heavy demands on expressivity and computability. This approach adopts many of the principles that underlie the ALEP formalism (Alshawi *et al.*, 1991) and has subsequently been put into practice in the ALEP grammars developed within the LSGRAM project.<sup>2</sup>

The PAS representation described in this paper clearly follows the view that at present computational grammars based on the typed feature-structures paradigm must comply to some sort of lean formalism in order to be useful from a practical point of view. That is to say the current state of the art in grammar building as it is developed within the HPSG community often involves too expressive mechanisms, which cannot be implemented in a satisfactory way in the computational environments at our disposal. In practice this means that a rather conservative approach is taken to some aspects (that is, to the most discussed and developed in the literature –the secondary predications, for example), whereas a somewhat adventurous approach is taken to some other (that is, to the less well-known aspects –the modifiers, for example).

The paper is organised as follows. In section 1 we discuss some general aspects which are relevant to justify the level of description adopted. Section 2 presents the general, basic options that we have adopted in the paper as a whole, and in particular in the exemplificatory sections. In section 3 we present the type system that we have adopted both in the kernel of the PAS and also in the collateral structures of the sign. Then follow sections 4 to 7 which provide the exemplification of the PAS for, respectively, the verbs, the adjectives, the nouns and the prepositions. Finally section 8 provides a tentative exemplification of a treatment of modifiers that is consistent with the rest of the paper.

---

<sup>1</sup>This project developed in 1993-95 and was financed by the Commission of the European Union. The following research centres took part in the project: IAI, Saarbrücken (Germany) (who acted as coordinator); University of Essex (Great Britain); UMIST, Manchester (Great Britain); University of Leuven (Belgium); Gruppo DIMA, Turin (Italy); and Pompeu Fabra University, Barcelona (Catalunya, Spain). A full description of the project, as well as of its reports, can be found in the URL address: <http://www.iai.uni-sb.de/REFMAN/home.html>. The final report of the project (Schmidt, 1996) contains a previous version of this paper.

<sup>2</sup>The LSGRAM project is also an EU funded project (LRE-61029) “Large-Scale Grammars for EC Languages”. It actually developed almost in parallel to the MLAP project, so that in many centres there were people writing linguistic specifications and people implementing them in real grammars. As a consequence there was a real interaction between the proposals made on paper and their realisation in a particular lingware.

# 1 Overview of the Coverage

## 1.1 Introduction

The aim of this section is to provide a preliminary description of the range of phenomena that are going to be treated in this paper and of some of the formal options taken in order to represent them.

The linguistic phenomena treated fall naturally into two broad classes, namely the structure of predicate and arguments, and the adjuncts. On the other hand, there are some more formal questions which also contribute to determining the scope of the work. Each of these aspects is treated in a subsection below.

## 1.2 The Structure of Predicate and Arguments

The structure of predicate and arguments is one of the aspects that have to be in the focus of the design of a purportedly general approach to the characterisation of the basic facts of language.

Predicates constitute one of the central backbones to the analysis of sentences in any language, since they are linguistic elements that organise the shape of many of the constituents that appear around them. This organising capacity is apparent from the syntactic point of view: each predicative element determines to a great extent the number and the form of its complements.

From a strictly monolingual point of view a syntactic characterisation of predicates is sufficient to provide a basis for the parsing of sentences: the differences in syntactic behaviour of predicates determines quite easily the grammaticality / ungrammaticality of sentences.

However if for some reason we are interested in a treatment that goes beyond the mere considerations of grammaticality, the syntactic properties of predicates are not enough. Whenever there should be an output to the processing that is comparable among the grammars of different languages, the syntactic characterisation of predicates must be overcome. This is obviously the case for MT systems, where the outcome of the analysis process must be comparable whichever is the language that has been analysed. But there are other applications for which such a comparable outcome is also desirable: natural language interfaces to data-bases, expert systems... And alternatively there might be an interest in having linguistic descriptions which are consistent among different languages in order, say, to reuse some of the modules for applications in other languages. In any of these cases an approach to the relation of predication which is restricted to syntactic properties is not adequate.

In an HPSG like framework this amounts to developing the content part of the sign in a way that yields characterisations of the constructions being analysed that are deeper (i.e., more abstract) than those that are obtained in the simple syntactic structure. To a



great extent then the purpose of this paper is to provide a description of the predicative relation that can be fitted into the content part of the predicative signs.

We are interested in providing a description to the predicative relation that is:

- broad enough to cover most phenomena related to predication in the major western languages
- deep enough to overcome some of the differences arising from the surface realisation in different languages
- specific enough to maintain in a strict way the link to the surface relation

Let us consider these aspects in turn. Firstly, the intended coverage should be broad enough. In particular we aim at characterising the relation of predication for all major categories in a consistent and coherent way. This means that all complementation patterns for verbs, nouns, adjectives and prepositions/adverbs are in the intended coverage of this section. In addition, it is intended that the treatment for all categories is consistent with one another; in a way that what is said about adjective complements is consistent with what is said about verb complements or that the treatment proposed for noun complements is not problematic for verb complements. To be more specific about these two examples, it is intended that the treatment of adjectives as predicates is compatible with what is said of raising and control verbs (and particularly of the verb *to be*), since in most occasions the predicative realisation of adjectives occurs when they are attributes (either of the subject or of the object); and what is said about the predicative use of adjectives should be compatible with what is said about their use as noun modifiers. With respect to the second example, it is clear that the discussion and proposal for the treatment of complements to nouns has consequences for the treatment of the complements of all categories; in this particular case, the difficulties in representing the arguments to nouns have been one of the fundamental reasons for adopting a particular basic structure of predicative signs for all categories.

Secondly, the coverage of the specifications should be deep enough. Thus, we aim at generality by intending to treat the basic phenomena for a wide range of European languages. Since there are obvious differences in the complementation patterns of these languages, this amounts to trying to provide descriptions of the predicative relation that are “neutral” with respect to these differences: thus the fact that the relation a complement has to its head is expressed by means of case, of a preposition or configurationally should not imply a different representation. Similarly the proposed system should provide similar representations for language specific different constructions (a complement to a noun should obtain the same basic representation whether it is in the genitive case or a prepositional phrase).

And finally, the description we are aiming at ought not to be as deep as possible, since depth in analysis alone does not provide consistency and coherence in the description. When trying to overcome surface differences one is tempted to go always deeper in the

analysis; this however, if unconstrained, is a bad move since the deeper one goes the more difficult is to maintain a levelness and coherence in the obtained descriptions. At the same time, the link between the abstract and the surface representations is to be maintained, if we do not want to introduce a large amount of indeterminacy in our descriptions. These are then the limits to the purported depth of description.

### 1.3 The Modifiers

All major categories can have adjuncts, whether they can have subcategorised complements or not. The abstract representation that we are aiming at should integrate the treatment of adjuncts in a way which is as general and coherent as possible.

There are two basic classes of adjuncts: those that modify predicative elements and those that modify non predicative (i.e., nominal elements). This is an important distinction, since the problems that are encountered when dealing with them are not the same. In HPSG there is now a reasonably established treatment for adjuncts to nominal signs, but there are not definite proposals for adjuncts to predicative signs yet. This means that the determination of a particular representation for adjuncts to predicative signs is one of the aims of this section.

Adjuncts to predicates can be further subdivided in those that can be interpreted in a restricted way (in a way which resembles most of the adjuncts to nominals) and those that cannot. Examples of the former class are locatives (whether temporal or spatial). Among the latter there are the adjuncts that can be interpreted as semantic operators (causatives...), and also those that are wholly external to the sentence (the constituents that do not clearly depend on the head).

An important aspect to be taken into account is that in many cases modifiers have the same form (and very similar meaning) than subcategorised complements. This of course has to be taken into account in order to produce results that are consistent with the treatment of complements. This is important with respect to, for example, adjectives and PPs. Many adjectives can either be predicates in copular constructions or noun modifiers, without any major change in meaning; consequently the treatment proposed should explain these similarities and try to produce related lexical entries for both cases. Similarly, there are PPs that can either be subcategorised for complements or adjuncts to a predicative sign; also in this case there is no substantial change in meaning from one another; the treatment of such adjuncts therefore should be compatible to that of the complements.

### 1.4 Formal Issues

From a formal point of view the aim of this section is to provide specifications for the design of the content part of the sign. This is going to take the form of a “lean” HPSG, since as described above (1.2), we are more interested in the coverage of our descriptions

than in their theoretical adequacy. In fact our proposal consists in a fairly detailed description of how the different predicate-argument structures (and the modifiers) can be represented in a consistent, levelled way. In some cases this is going to provide specifications which are clearly behind what has been already obtained in the HPSG literature (like in the case of some adjuncts to predicative signs, as in the proposal by (Kasper, 1994), where a more semantically oriented proposal is made for some of these modifiers), but in some others our proposal is going to treat phenomena that are not covered by current HPSG literature (like in the case of the arguments to predicative nouns or many complement types treated in the section on the lexical representation of verbs). To this end, we propose in section 3 a type system which covers all of the cases treated; it basically follows (Pollard & Sag, 1994) typing, but it obviously deviates from it in some ways to permit the specific proposals contained in this section.

It is also important to bear in mind that the specifications provided are to be considered as the input to parsing. In the exemplification performed in the final sections of this paper, the predicate-argument structures (and the modification relations) are represented in lexical entries of the relevant categories; this however has not to be understood as a proposal for the organisation of the lexical material. In this paper, we are viewing lexical entries just as they enter parsing: we are not concerned with the way in which they are created or stored. We are thus assuming mechanisms for the lexical organisation such as lexical rules or a hierarchical organisation that would provide the necessary simplification to the amount of lexical entries we are assuming.<sup>3</sup> An important question is the extent to which lexical generalisations are dependent on the use of the Lexical Redundancy Rules mechanism, that is to say, to what extent they can only be expressed by means of LRRs. There are reasons (basically related to expressivity and restrictiveness) to try to limit the use of LRRs, and to employ underspecification in the type system instead; this however puts very definite requirements in the expressivity of the formal system. In Markantonatou & Sadler (1996; to appear) there is a discussion of this subject as well as a very thorough exploration of the limits imposed by the formal assumptions adopted in this work (as expressed by P. Schmidt in the chapter on the formal assumptions of the final report of the project (Schmidt, 1996)). As shown there in some cases underspecification can be used to express the relation between different lexical entries, but there are also cases in which this is not possible, so that the only formal means to express them are LRRs. In the rest of this paper, we are not going to discuss the formal means of expressing lexical generalisations; instead whenever a generalisation is foreseen, we just indicate it explicitly (either by the mention of a possible LRR, or by a more general statement).

---

<sup>3</sup>An interesting discussion of these questions can be found in the chapter on lexical generalisations of the final report of the project (Schmidt, 1996).

## 2 Basic Options

### 2.1 Introduction

Although it is generally agreed that surface syntactic relations are not enough to characterise the predicate-arguments relations, there is no commonly accepted way of characterising the deeper relations that should be substituted for the former. It can be easily observed that in most cases the formal description does not go beyond the simple labelling of arguments and relations in order to show to what extent this description differs from the surface one.

However in the research on transfer based MT a tradition has developed of establishing a particular level of description which is rather fit, it is claimed, for stating the more abstract level of representation upon which the transfer operation is performed. This is a work started within GETA (on the transfer based MT system Ariane (Vauquois, 1988)), but which has been explored at length in Eurotra. The result was a description of a level of description (called **IS** – *Interface Structure*) which could be used as the basis of the transfer operation (Allegranza, et al, 1991; Durand et al, 1991; Badia, 1992).<sup>4</sup>

The basic insight of this work was to determine a level of description which is in between the surface syntactic and the full semantic, and which corresponds to what might be called *deep syntax*. To a certain extent the resulting description was a hybrid one which derived from the results obtained in a broad class of theories and frameworks. The importance given by LFG to grammatical functions (Bresnan, 1982), the deep grammatical functions of Relational Grammar (Perlmutter, 1983; Perlmutter & Rosen, 1984), the notion of head derivable from X-bar theory (Jackendoff, 1977), the deep structure of GB (Chomsky, 1981) and the thematic relations as defined in this theory (Jackendoff, 1983; Wilkins, 1988), the recent applications of Valenz Grammatik to MT (Somers, 1986; Somers, 1987; Kirschner, 1987; Schubert, 1987; Gebruers, 1988), the transitivity structure of sentences as defined in Systemic Grammar (Halliday, 1967-8; Halliday, 1970), provide all the background upon which the Eurotra Interface Structure was defined.

The work on MT we are referring to was developed within stratificational approaches to analysis. Consequently the level of description obtained was really conceived as a grammar which could build a particular structure fulfilling a particular set of requirements. Let us consider it in some detail.

The interface representations in such an approach naturally contained all the information that was needed to perform the transfer operation; and it must be borne in mind that most of this information was only present in the representation in the form of features that decorated the nodes of the tree representations. What interests us here however is the elements in the representation that structured it, because these were the predicate-

---

<sup>4</sup>Much of the work on the basic structure of representations reported in these references, and also much of what follows here, is clearly derived from work by Lee Humphreys who led the research on predicate-argument structure within Eurotra.

argument relations. That is to say, the structuring of the interface representations was not based on constituency, but on the predicate-argument (and modification) relations that the elements of sentences established with one another. A real dependency grammar then determined the construction of such structures.

The basic units of the grammar were dependency constructions, consisting of a single governor and zero or more non governor dependents (which were also dependency constructions, thus recursively defining indefinitely deep constructions).

Governors were characterised by the factors standardly assumed in the literature (Zwicky, 1985; Hudson, 1987):

- the governor determines the type of the denotation of the whole construction,
- the governor determines the agreement features of (some of) its dependents,
- the distribution of the governor alone equals the distribution of the governor plus their dependents, and
- the governor bears a frame which determines most of the properties of its dependents.

On the other hand, dependents were characterised by the converse factors:

- the dependents do not determine the type of the denotation of the whole construction,
- the dependents are subject to the agreement features imposed upon them by the governor, and
- the distribution of the dependents need not to coincide with that of the whole construction.

For some dependents the following, additional condition also held:

- the dependents have to fulfill the conditions imposed upon them by the frame of the governor.

Thus two different sorts of dependents were distinguished according to whether the last condition (on being restricted by the frame of their governor) applies to them or not:

- **arguments**: dependents which enter into the relation expressed by the governor.
- **modifiers**: dependents that do not enter into the relation expressed by the governor.

As corollaries of, or further restrictions on, the preceding definitions, the following was stated:

- all governors are lexical, or else are empty nodes that are coindexed with lexical nodes.<sup>5</sup>
- a construction has one and only one governor
- all dependents correspond to a semantic relation
- the restrictions the governor imposes upon their dependents are locally specified.

In order to ensure that the correct projections were formulated between the representations of different languages, a labelling was applied to all dependency relations (see section 2.3 below).

Of course such a characterisation cannot be directly applied to a TFS system and needs some adaptation. There are however aspects of it that can be useful in determining a third solution, in addition to the two immediate ones: to use a strictly syntactic (probably functional) characterisation of the relation between predicates and their arguments, or to move to a fully semantic representation (either adopting a standard logical form, or applying some semantic theory –such as *situation semantics*). This is highly wellcome since it is not at all clear that either of these two solutions provides a reasonable characterisation for determining the shape of the content of linguistic signs in TFS formalisms.

The adoption of a strictly syntactic description (even if it is functional in nature) is totally linked to the surface structure of sentences. Consequently this is a level of description which is not appropriate for the work in a multilingual context in a modular way. Even for the restricted set of languages we are having in mind (namely, the western European ones), there is no common set of syntactic functions, and more importantly there are lots of discrepancies in the specification of the argument structure of corresponding (translationally equivalent) predicates. Thus what in one language is a direct object in another one might be a prepositional object; or a full sentential complement might be a predicative complement (xCOMP) in another. For examples of these discrepancies see the section on the detailed description of predicate-argument structures of verbs (section 4).

Although for strictly monolingual settings such an approach might be useful, it is wholly inadequate for multilingual ones, even if we are not thinking directly in MT. It is convenient in such a setting that the abstract representations of sentences in the different languages considered are as much similar as possible. If we think of applications in which a module is introduced that is independent from language (information retrieval, interface to data bases...) the more similar are the representations from different languages the better since the link between the linguistic modules and the external one is going to be much simpler (there will not be major differences for the different languages).

On the other hand, the fully semantic representation is problematic in that there is no general approach that can be applied to a large range of linguistic phenomena. There are

---

<sup>5</sup>Note that this condition was imposed because the structures being built were trees (and not, say, graphs), so that proper structure sharing was not permitted.

of course proposals to use thematic roles to characterise the relations that the predicates have with their arguments; there is quite a large literature on cases, semantic roles or thematic roles. However there is no general treatment of such abstract relations that is both general enough to deal with all major complementation patterns in languages and consistent with an overall semantic perspective.

A characterisation of the predicate-argument relations that is both strongly dependent on syntactic structure and abstract enough to overcome the major differences between equivalent surface structures of sentences in the different languages can be obtained from the efforts mentioned above about the definition of a transfer interface for MT.

Given the basic assumptions of the MLAP project (Schmidt, 1996:ch.2), any such theory needs to be adapted to a TFS format, in a way that is consistent with the basic principles of a lean version of HPSG. There are two dimensions upon which this conversion has to be performed: the labelling of predicative types and arguments, and the basic structure of the content types in general. These are indeed two aspects of HPSG which raise problems when trying to widen the coverage of the grammar.

In standard HPSG there is no consistent labelling of predication types nor of argument classes. Instead the relation represented by a predicate is used as the label of the type (actually, a subtype of the type *qfpsoa*) and arguments are labelled with relation-related tags (e.g., *liker* and *liked* for the relation type *like*). Indeed such a treatment is not adequate to express generalisations about the predicate-arguments structure.

On the other hand, there is no treatment of modifiers to predications in standard HPSG. Although the proposal for modifiers to nominal objects conforms to the standard approaches to modifiers in semantically oriented theories, the application of this proposal to predicative signs is not straightforward. Indeed the same problems occur that are found in any logical form treatment of modifiers to predication.

These two aspects then have to be restated in our TFS lean formalism, since there are no clear proposals for them in HPSG proper. To these two questions we turn our attention in the next two sections.

## 2.2 The Basic Structure of the Content of Signs

There are two fundamental problems concerning the generalisation of the structure of the type *content* (as described in Pollard & Sag, 1987; 1994) to a large range of phenomena. These are the representation of the PAS of nominals and the integration of adjuncts to predicates with their PAS. Let us start by considering the problem of the representation of adjuncts to predicative signs.

The fact that adjuncts to predicative signs do not fit in the content type of predicative signs basically derives from the lack of an index to which adjuncts could impose restrictions.

This problem can be solved by introducing an index for eventualities in a parallel way as

nominal signs introduce an index over individuals. In this way the FS for the semantics of predicative signs has a new attribute INDEX whose value is a variable of a new type, namely that of an eventuality. This move implies the introduction of an additional event argument in the PAS, whose value is structure-shared with that of the index. See for example the FS for a verb like *sing*:

$$(1) \left[ \begin{array}{l} \text{INDEX} \quad \boxed{0} \text{ eve} \\ \text{RESTRICTION} \quad \left\{ \begin{array}{l} \text{REL} \quad \textit{sing} \\ \text{ARG-EVE} \quad \boxed{0} \\ \text{ARG1} \quad \boxed{1} \end{array} \right\} \end{array} \right]$$

Note that in this way we have a natural place where to insert the content of the adjuncts; adjuncts then can impose additional restrictions to the eventive index thus introduced.<sup>6</sup> In its simplest form the resulting structure is like the following:

$$(2) \left[ \begin{array}{l} \text{INDEX} \quad \boxed{0} \text{ eve} \\ \text{RESTRICTION} \quad \left\{ \begin{array}{l} \text{REL} \quad \textit{sing} \\ \text{ARG-EVE} \quad \boxed{0} \\ \text{ARG1} \quad \boxed{1} \end{array} \right\} \& \textit{psoa}[\ ] \& \textit{psoa}[\ ] \& \dots \end{array} \right]$$

This FS resembles completely the one corresponding to the content of nominal signs, as presented in Pollard & Sag (1987; 1994). In these FS however the first restriction to which the index is submitted is the state of affairs denoted by the head and its arguments (if any), so that we have the *psoa* corresponding to the basic predicate and the possible *psoa*'s corresponding to the adjuncts in the same set. This could create some difficulties since there is no mechanism which compels to look for an argument in the basic predicate *psoa*, i.e. the type system could not fulfill its restrictive function anymore. Particularly this could create difficulties in generation: consider the FSs corresponding to the value of the content of phrasal signs in which a couple of restrictions appear which are not distinguishable from one another in a principled way, such as those formed by a noun and an adjective or by a noun-noun compound. An example of such a structure could be the following (corresponding to the phrase *red house*):

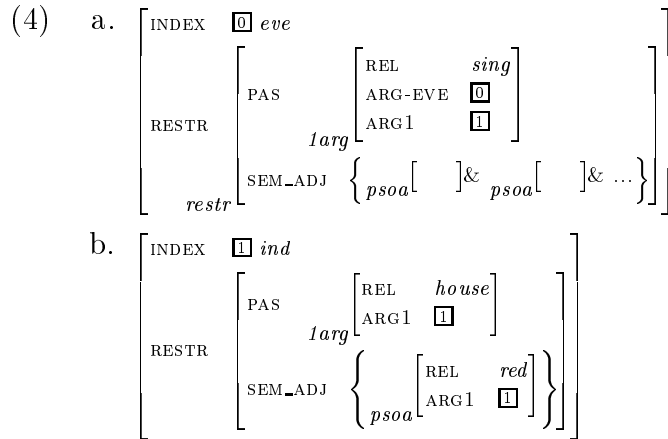
$$(3) \left[ \begin{array}{l} \text{INDEX} \quad \boxed{1} \text{ ind} \\ \text{RESTRICTION} \quad \left\{ \begin{array}{l} \text{REL} \quad \textit{house} \\ \text{ARG1} \quad \boxed{1} \end{array} \right\} \& \begin{array}{l} \text{REL} \quad \textit{red} \\ \text{ARG1} \quad \boxed{1} \end{array} \end{array} \right]$$

To generate a syntactic structure from this content a decision is to be taken as to what is the head of the construction. It seems therefore to be more appropriate to have the *psoa* provided by the head (and its arguments, if any) and those introduced by the adjuncts as values of a different attribute. Thus the value of RESTR is of type *restr* and two attributes are defined for it: PAS which takes the *psoa* provided by the head as value, and SEM\_ADJ

<sup>6</sup>This is, of course, the proposal initially made by Davidson (1967), which is now known as eventive semantics.



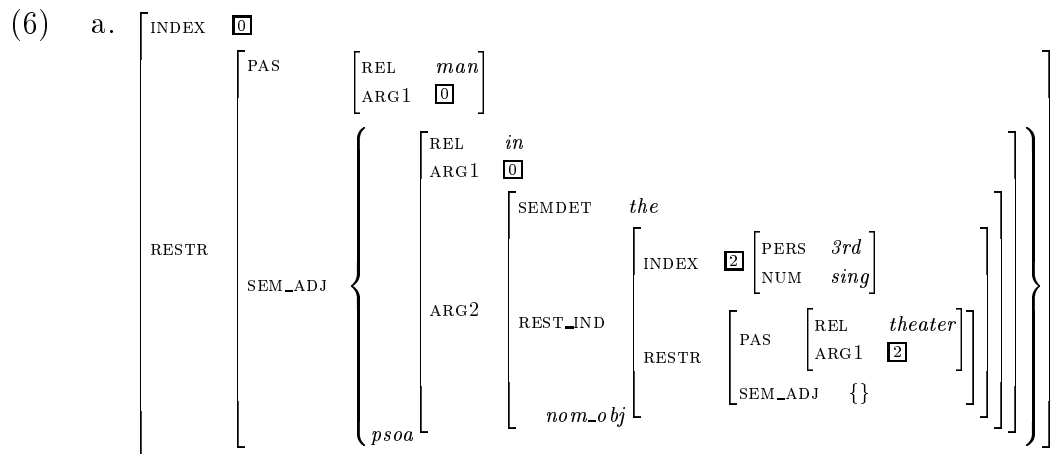
which takes as value a set of contents (namely, those provided by the adjuncts to the head). See for example the FS for the same verb as in (2) and of the same nominal phrase as in (3):

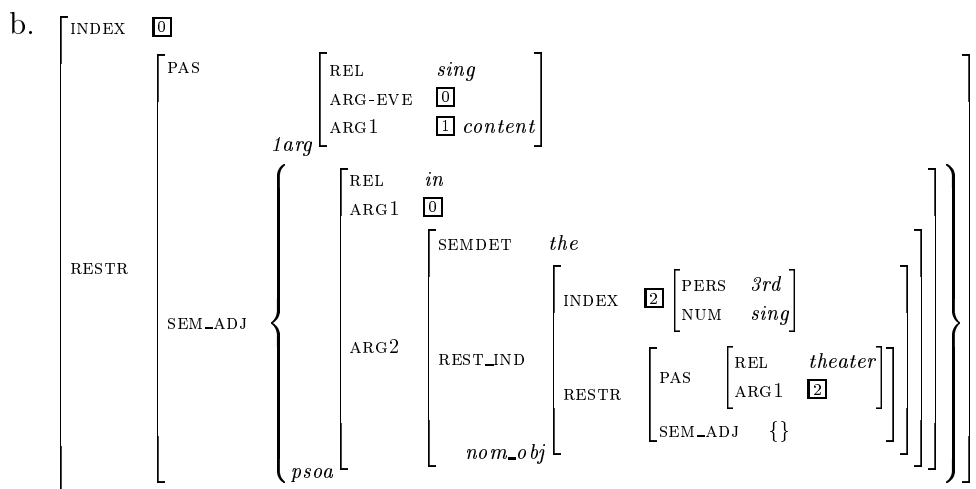


Note that this move has been taken in a parallel way for predicates and for nouns so that we obtain a similar representation for both verbal and nominal signs. Such a FS also provides a uniform representation of those constituents that can be both adjuncts of a nominal and of a predicate. Thus a PP like *in the theater* for example has the same representation in the following two contexts:

- (5) a. man in the theater  
 b. sings in the theater

In both cases the adjunct is linked to an index, namely the index of the individual and the index of the eventuality respectively, which it restricts.



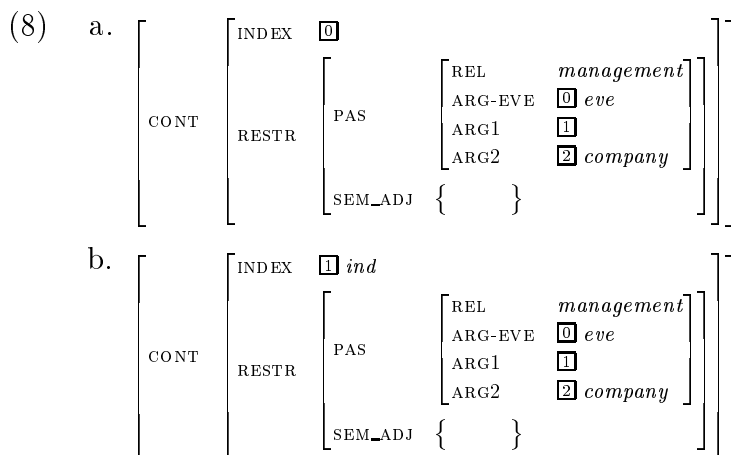


An additional benefit of the approach just adopted is that noun arguments can easily be represented. The introduction of an eventive variable brings together the formal structure of nominal and verbal signs, as has just been shown above.

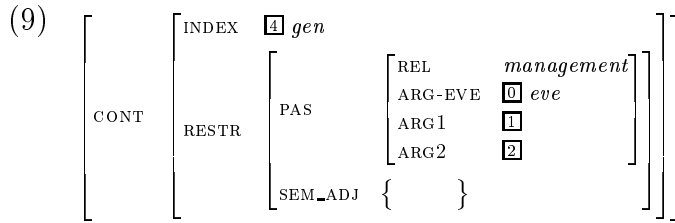
Were we to accept the definition of types with feature structures, the distinction between functional and dynamic deverbal nouns could be treated by assuming a common “underspecified” FS with a generic index for all predicative nominal signs. Therefore, the distinction between both would become just a difference about the kind of entities that are denoted by them. Consider the following examples:

- (7) a. The management of the company is difficult  
 b. The management of the company has decided...

In (7a) we have a dynamic nominalisation that introduces an index over eventualities, whereas in (7b) the subject is a functional nominalisation which therefore introduces an index over entities. See the following feature structures for these two possible readings (respectively):



These two FSs can be subsumed by a more general, underspecified one like the following:



Note that this underspecified FS can be taken as a supertype of the other two.

### 2.3 The Labelling of Predicative Types and Arguments

In the exemplification sections of this paper we adopt a very simple labelling of predicative types and arguments: we follow standard practice in maintaining two semantically unrestricted arguments, and introducing mild semantic labels for the rest. Thus the first and second arguments are labelled in a semantically unspecified way (simply, ARG1 and ARG2). The rest of the possible arguments are labelled in a relatively transparent way: ARG\_2P for the second animate participant; ARG\_2E for the second inanimate participant; ARG\_PE for the perceiver in raising predicates; ARG\_AS for attributes of the subject; ARG\_AO for attributes of the object; and ARG\_ADJ for arguments with a semantics which resembles that of adjuncts. The latter class of arguments are further specified as ARG\_PLACE, ARG\_GOAL, ARG\_ORIG, (for the locational arguments), ARG\_MEASURE (for the measure argument of verbs like *sell* and *buy*) and ARG\_ASSOC (for a particular class of “associated” arguments –see subsection 4.2.4.2 below). And predicates are typed for the list of arguments that they ask for (e.g., *arg\_1\_2*, *arg\_1\_2\_2P*, *arg\_1\_PLACE*, and so on).

### 2.4 Classes of Complements

In this subsection we classify the different classes of syntactic constituents that can appear as arguments, and we discuss the representation of some of them.

#### 2.4.1 Noun Phrases

The first, possibly most common complement class is the NP. Almost all verbs allow for NP complements in some of their subcategorised for positions. Yet not all NPs have the same semantic value as they certainly do not have all the same syntactic structure: there are referential and quantified NPs; and determinerless nouns; and clearly predicative NP constructions. However we are not going to discuss here the different representations for them (see section 6 for some discussion on the representation of predicative nominal signs, and the chapter on determination of Schmidt (1996), for a general discussion on nominal signs). Let us here only mention that NPs with a referential (or quantified) import are going to be represented as possibly quantified indexes the reference of which

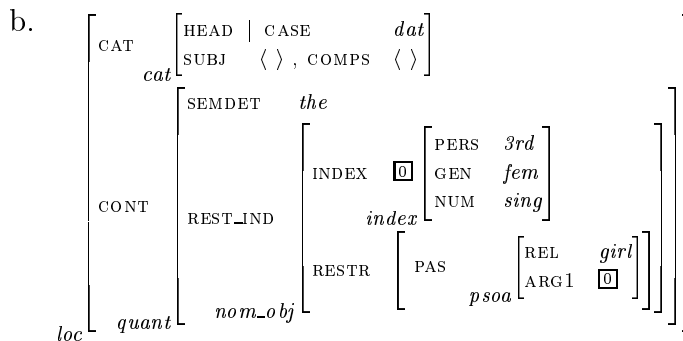


is semantically significant. Consequently the preposition can be considered the head of the construction, since it determines its syntactic and semantic properties; in general, the semantic value of these PPs depends to a great extent on the semantic value of the preposition, and usually it is similar to that of some adverbial phrases. In a natural way all prepositions possibly occurring in such a construction share some semantic value.

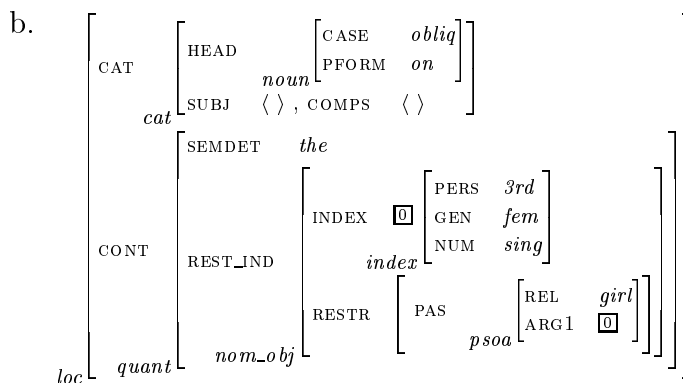
Given the differences just mentioned it is natural that these two sorts of PPs are given different representations. PPs with strongly bound prepositions are represented as NPs, with some additional information, and PPs with weakly bound prepositions are represented as PPs (and in a very similar way to ADVPS –in fact, we do not establish any principled distinction between these two classes of constituents).

Thus a PP with a strongly bound preposition is going to be given a representation which is basically nominal, but which has been enriched by some head information concerning *case* and *pform*. In the following sections on the representation of the lexical signs it is assumed that some PPs with strongly bound prepositions only result in case information, whereas some others yield an underspecified case (e.g., *obliq*) which is further specified by the feature PFORM. Note that this is only a rough approximation which cannot be fully developed unless all the particular facts are examined about strongly bound prepositions and case in the relevant language. However, in order to properly exemplify the different complementation possibilities representations as the following for the *local* type are assumed throughout this paper:

(11) a. to the girl



(12) a. (depend) on the girl



Note that this representation can be obtained by a rule which is similar to the ID schema 4 in Pollard & Sag, 1994, in which the head daughter is the NP and the marker daughter is the preposition.<sup>7</sup>

Note that the treatment proposed here for strongly bound prepositions implies a lexicalist approach to unbounded dependencies, especially for languages like English which allow preposition stranding. In these cases the basic predicate has to be lexically related to a predicate having a preposition as a particle, thus being treated like a phrasal verb (see subsection 4.2.3.5). Alternatively, strongly bound prepositions could be treated in these languages as syntactic heads that do not stand for any semantic relation.

On the other hand, PPs with weakly bound prepositions are given representations in which the semantic value of the preposition is present. This is obviously obtained by treating the preposition as the head of the construction and including the content of the subcategorised NP as the value of their ARG2 feature. Note also that all PPs headed by semantically full prepositions have an ARG1 as well. Let us examine briefly the three possible cases.

Firstly, these PPs can occupy ARG\_ADJ positions. In these cases they introduce an index, which is of type *locative*, and their ARG1 is coindexed with that index. Thus the preposition can be said to formulate a relation between the denotation of the subcategorised NP and a particular location (i.e., the one denoted by the whole PP).

Secondly, weakly bound PPs may appear in attribute position. In these cases their ARG1 is either occupied by the content of the raised NP (in the case of attributes to the subject), or is structure shared with the content of the controller NP (in the case of the attributes to the object).

Finally, semantically full PPs can be modifiers (to nouns or to verbs). In these cases, the ARG1 is structure shared with the index introduced by the head of the construction.

Thus the general shape of the *psoa* type for a PP with a semantically full preposition is going to be as follows (exemplified for preposition *in*):

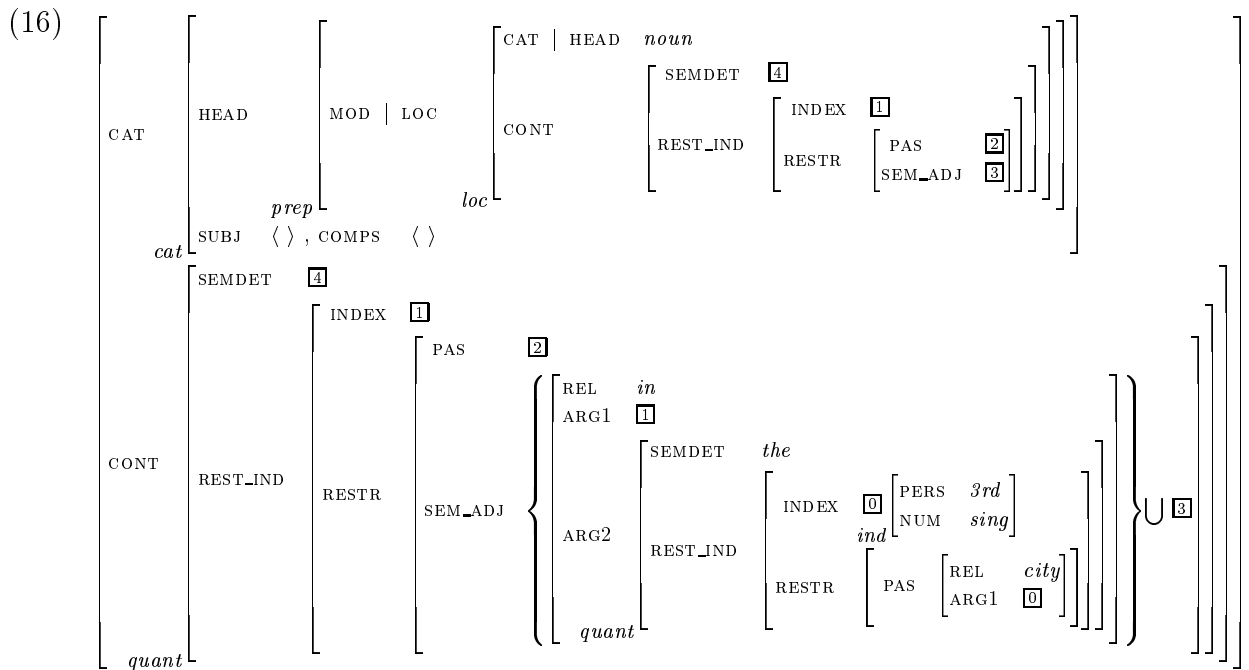
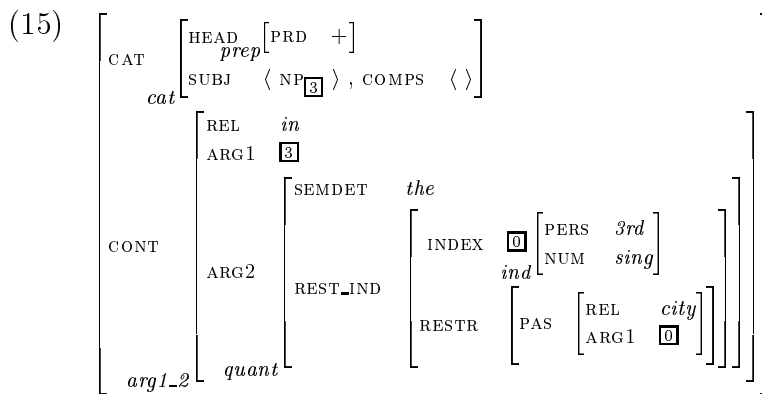
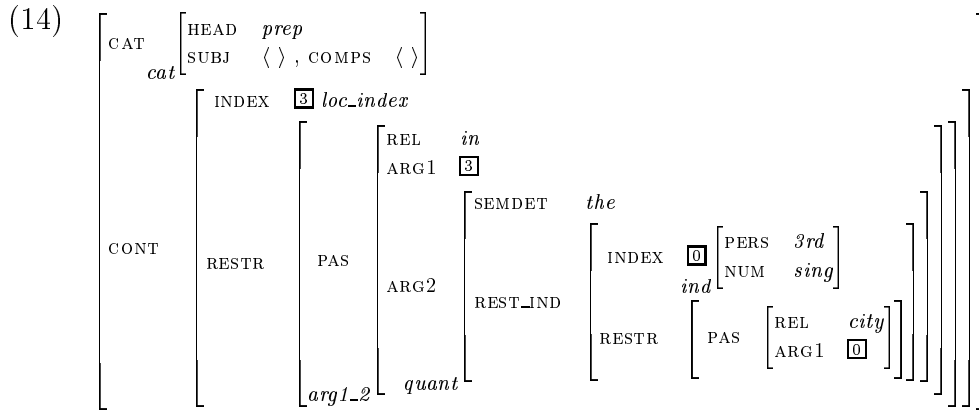
$$(13) \quad \underset{arg1\_2}{\left[ \begin{array}{ll} \text{REL} & in \\ \text{ARG1} & \textit{“the external argument”} \\ \text{ARG2} & \textit{“the content of the NP”} \end{array} \right]}$$

The PP *in the city* can be an ARG\_ADJ (as in (14)), an attribute (as in (15)), or a modifier (as in (16)).

---

<sup>7</sup>However, a very similar effect can be produced by treating the preposition as the head of the construction (albeit a semantically empty head), provided that the content of the complement of the preposition is unified to the content of the PP itself. In this case, no syntactic distinction is established between prepositions, and all PPs have a *prep* head and are built by ID schema 2.

# The Predicate-Argument Structure



### 2.4.3 Other Predicative Complements

Two different kinds of constituents other than PPs can occur in predicative subcategorised for constructions: some APs and NPs, as they occur in the attribute (of the subject and of the object) positions. Two characteristics in their representation mark this type of constituent: a head feature PRD + and a member in the SUBJ list (i.e., they are not “saturated signs”). The PRD feature indicates that this constituent has to occur in the contexts just mentioned. On the other hand, for the attributes of the subject the element in the SUBJ list, which can be considered the subject or the external argument, is necessary so that the subject of the main sentence is assigned a  $\theta$ -role argument position, since the subject of copular verbs are not assigned any (see section 4.2.3.3 for the treatment of copular verbs as raising verbs). In a similar way, the object of predicates having an attribute of the object establishes a double  $\theta$ -relation: with the main predicate and with the PRD + constituent. Thus the following is a simplified lexical representation for the *loc* type of an adjective occurring in attribute position (*ill*):

$$(17) \quad \left[ \begin{array}{l} \text{CAT} \\ \text{cat} \\ \text{CONT} \\ \text{arg1} \end{array} \left[ \begin{array}{l} \text{HEAD } \textit{adj} \left[ \text{PRD } + \right] \\ \text{SUBJ } \langle \text{NP } \boxed{0} \rangle \\ \text{REL } \textit{ill} \\ \text{ARG1 } \boxed{0} \end{array} \right] \right]$$

### 2.4.4 Sentential Complements

Finally there are the sentential complements, which can be either finite or infinite. Finite sentences are introduced by a marker (e.g., *that* or *whether*) that enables them to appear in complement positions. This marker functions in a way similar to that of the strongly bound prepositions: it takes well-formed, complete finite sentences and prepares them to appear in dependent positions; and consequently it is treated similarly. The marker and the finite sentence are the only two daughters of this type of head-marker construction, in which the sentence is the head and the marker just adds a little information onto its feature structure: basically the category feature that expresses that the sentence has been marked by a particular marker (e.g., MARKING *that*).

On the other hand, infinite sentences that appear in subcategorised for positions have in a natural way an element in their subcategorisation lists (in particular, in their SUBJ list) which has not been cancelled and which is the one that either is controlled by another argument of the main sentence or has been raised onto a complement position of the main predicate (see section 4.3, for the corresponding representations).

## 2.5 The Syntactic Representation of Complements

There is no question here to study the different surface structures that the different PAS may have (see the chapter on Phrase Structure of Schmidt (1996), where this problem



has been dealt with). There is however something to be said on this respect, since in some exemplifications we need to refer to the surface structure of the phrases that result in the PAS we are discussing.

In the rest of this paper we are assuming that the different constituents of the phrases we are discussing appear in the surface order which might be called canonical. Thus we assume that verbal complements appear in the surface in exactly the order that is expressed in the obliquity hierarchy of the subcategorisation lists. This is clearly an idealisation which cannot be assumed in general in the construction of realistic grammars; however the kind of problems that relate to surface order are of a different nature to those relating to the representation of PAS. In particular we assume that some mechanism is available, similar to the ones used in the grammars developed for the ET10/52 project (Markantonatou & Sadler, 1994a). At least a couple of grammars are presented there that simulate in an ALEP environment lexical rules that perform the permutation (Schmidt, 1994; Badia, 1994). However the reader must be conscious of the difficulties of this treatment, in particular with respect to the ambiguity in parsing that may be generated by such rules and the strategies used in the grammars just mentioned (Markantonatou & Sadler, 1994b).

There is an aspect however that bears some relation to the representation of surface structure, which seems worth being discussed briefly now. In standard HPSG (Pollard & Sag, 1987; 1994) there is a single list containing all subcategorised for complements of the verbal predicate, including the subject or external argument. However in the final chapter of Pollard & Sag (1994), following work by Borsley (1989; 1990), the authors suggest that the SUBCAT list should be divided into two separate lists: one including the subject and the other one for the rest of the complements.

The main reasons for adopting the two separate lists (i.e., the COMPS list and the SUBJ one) are that VPs are basic constituents of sentences (as it is recognised in most theories) and that the subject position cannot be equated to the other argumental ones: it can be filled by highly different sorts of constituents and the behaviour of traces and extractions is different from other sorts of argument positions. On the other hand the work in the ET10/52 project has shown that the treatment of the subject (either in a separate list, or as the first element in the common SUBCAT list) interferes with the possibilities of dealing with the surface order variations of the complements in an efficient way (Markantonatou & Sadler, 1994b).

In addition, it must be borne in mind that a decision with respect to this topic must take into account as well the specificities of the language being described. For example, the fact that Spanish or Italian are *pro-drop* languages must be taken into account when trying to develop a grammar for any of them; thus the interaction between *pro-drop* and the different word orders in a clause (especially the appearance of the subject in post-verbal position and that of the object —and many other complements— in pre-verbal one) has to be taken into account when designing the parsing strategy and the structure of the subcategorisation lists. Similarly, when trying to develop a grammar for Dutch the phenomenon known as the Dutch crossing dependencies has to be taken into account;

here again there seems to be an interaction between the use of binary branching and the distribution of the complements in two separate lists.

In the exemplification sections of this paper we have chosen to write all our lexical entries with a divided subcategorisation list, so that both the SUBJ and the COMPS lists appear.

## 2.6 The Parallelism between the Syntactic and the Semantic Representation of Complements

When trying to devise lexical entries for predicative elements, the linguist tends to see the syntactic and the semantic information in parallel; that is to say, there is a natural tendency to consider at the same time the subcategorisation information and the predicate argument structure of predicates, in a way that formulates explicitly that there is a relation between the arity of the subcategorisation lists and the arity in the predicate argument structure. That such a relation exists is a fact that no one would be willing to deny: mono-valent verbs, for example, have a unique element in their subcategorisation lists and also have only one argument (besides, of course, the event-argument); di-valent verbs have a couple of complements in their subcategorisation list and have two distinct arguments in the content part of the sign; and so on.<sup>8</sup>

It is also an important benefit of lexicalist, monostratal formalisms (like the one we assumed in the project where this paper developed) that it is possible to state this relation in arity in a single level of description. Thus no projection mechanisms are needed between levels and consequently the need for different strata in the grammar disappears as well. The question however remains whether this relation is somehow restricted by the tools at our disposal, so that the theory (or formalism) adopted contains some explanatory power about the way in which these relations are established in languages.

Of course a first possibility is that one relation is derived from the other. In some recent work in theoretical linguistics (particularly within the LFG framework) there is an attempt to deriving surface relations (possibly expressed in terms of grammatical functions) from some more abstract, deeper relations. In our terminology, it could be said that they attempt to derive the subcategorisation lists from the argument structure of predicates. A proposal was devised within the MLAP project that is at the base of this paper to apply this approach to a typed feature structure formalism like ours (Markantonatou & Sadler, 1996; to appear). Of course an additional benefit of proposals like these is that they not only (try to) explain surface grammatical in terms of the deeper, more semantically oriented, argument structure, but that by the same token they also explain (most of) the alternations the different predicate types can undergo. When applied to monostratal typed feature structure formalisms, such an approach however presupposes a quite strong machinery and a powerful descriptive semantics, both of which are outside the scope of the formal limits of our work. Thus we are left with

---

<sup>8</sup>Of course, there are cases in which this correspondence is not a one-to-one relation (as with absent subjects in *pro-drop* languages, or with control predicates).

the need to trying to be as restrictive as possible in the formulation of the relationships between subcategorisation and argument structure.

The first possibility that comes to the mind consists in formulating this relationship in the type system. That is to say, to write the type system in such a way that the impossible correlations in the language are not permitted by the type system. This is of course possible, but it implies that types are not defined in terms of appropriateness conditions but with feature structures (or, equivalently) paths. See the discussion of both approaches to type system writing, and of the consequences this implies, in the basic decisions chapter of Schmidt (1996).

A second possibility consists in implementing the parallelism in a macro system that would simplify the task of writing lexical entries for the grammar. That amounts to declaring previously to the beginning of the writing of lexical entries the permissible correlations between subcategorisation and argument structure. To a great extent this would provide the lexicographer with the same facilities a restrictive type system would; however it is not going to forbid the writing of lexical entries which are extraneous to the system. Of course this is so because there is no compilation process upon the macros. Nonetheless if a careful use of this approach is made, it can simulate rather well in our lean system what in more powerful systems would be performed by the type hierarchy.

Finally, the third possibility consists in formulating the relationships in just the lexical entries, in a similar way to what is been done in the exemplification sections of this paper. This is of course the less restrictive approach, since there is nothing that would prevent the grammar writer to create inadequate lexical entries, which would correlate any subcategorisation list to any argument structure.

## **3 Proposal for Formalisation: The Type System**

### **3.1 Introduction**

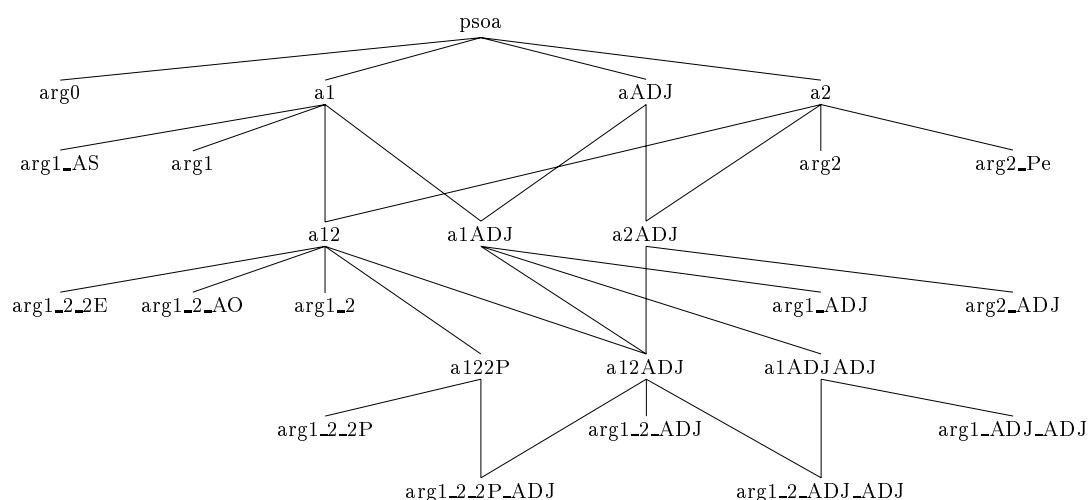
In this section we show first the type system that underlies the PAS structures in the exemplification of the following sections, and secondly a variant of the HPSG type system (Pollard & Sag, 1994) which has been adopted to provide the exemplification below.

### **3.2 A Proposal for the PAS Type Hierarchy**

One of the most interesting properties that the type hierarchy should have is that it allows as much type inference as possible. This is so because the analysis process is much eased if many of the relations that are relevant to that process are derived directly from the type structure. In the following table we provide a small fragment of a possible type declaration for the PAS chosen that conforms to this desire.

TYPE	ATTRIBUTES	SUBTYPES
<i>psoa</i>		<i>arg0, a1, a2, aADJ</i>
<i>arg0</i>		
<i>a1</i>	ARG1 = <i>content</i>	<i>arg1, arg1_AS, a12, a1ADJ</i>
<i>arg1</i>		
<i>arg1_AS</i>	ARG_AS = <i>content</i>	
<i>aADJ</i>	ARG_ADJ = <i>content</i>	<i>a1ADJ, a2ADJ</i>
<i>a2</i>	ARG2 = <i>content</i>	<i>arg2, arg2_Pe, a12, a2ADJ</i>
<i>arg2</i>		
<i>arg2_Pe</i>	ARG_PE = <i>content</i>	
<i>a12</i>		<i>arg1_2, arg1_2_AO, arg1_2_2E, a122P, a12ADJ</i>
<i>arg1_2</i>		
<i>arg1_2_AO</i>	ARG_AO = <i>content</i>	
<i>arg1_2_2E</i>	ARG_2E = <i>content</i>	
<i>a122P</i>	ARG_2P = <i>content</i>	<i>arg1_2_2P, arg1_2_2P_ADJ</i>
<i>arg1_2_2P</i>		
<i>arg1_2_2P_ADJ</i>		
<i>a1ADJ</i>		<i>a12ADJ, arg1_ADJ, a1ADJADJ</i>
<i>arg1_ADJ</i>		
<i>a1ADJADJ</i>	ARG_ADJ2 = <i>content</i>	<i>arg1_ADJ_ADJ, arg1_2_ADJ_ADJ</i>
<i>arg1_ADJ_ADJ</i>		
<i>a2ADJ</i>		<i>a12ADJ, arg2_ADJ</i>
<i>arg2_ADJ</i>		
<i>a12ADJ</i>		<i>arg1_2_ADJ, arg1_2_2P_ADJ, arg1_2_ADJ_ADJ</i>
<i>arg1_2_ADJ</i>		
<i>arg1_2_ADJ_ADJ</i>		

Note that this type declaration has been simplified basically in the sense that no distinction has been established between the different sorts of ARG\_ADJs. Note as well that this type hierarchy conforms to the type specification hierarchies of Carpenter (1992), and therefore the expressive and inferential power discussed in Carpenter’s book applies directly to it. The inferential power of such a hierarchy is best seen in a diagram showing the partial order between the types:



### 3.3 The Type System Used in the Exemplification

In this section we present the type system that underlies the exemplifications that are provided in this paper. It should be noted that this type system cannot be seen as settling the different aspects that are mentioned in it. Except for some relevant questions that we discuss below, it is only intended to help the reader to understand the examples that follow in the different sections.

TYPE	ATTRIBUTES	SUBTYPES
<i>sign</i>	PHON = <i>list of phon-strings</i> SYNSEM = <i>synsem</i>	
<i>synsem</i>	LOC = <i>local</i>	
<i>local</i>	CAT = <i>category</i> CONT = <i>content</i>	
<i>category</i>	HEAD = <i>head</i> SUBJ = <i>list of synsems</i> COMP = <i>list of synsems</i> SPR = <i>list of synsems</i> MARKING = <i>marking</i>	
<i>head</i>		<i>substantive, functional</i>
<i>substantive</i>	MOD = <i>modsynsem</i> PRD = +,-	<i>noun, verb, adj, prep</i>
<i>modsynsem</i>		<i>synsem, none</i>
<i>noun</i>	CASE = <i>atom</i> PFORM = <i>atom</i>	
<i>verb</i>	VFORM = <i>atom</i> AUX = +,-	
<i>adj</i>	PRD = +,-	
<i>prep</i>	PRD = +,-	
<i>content</i>		<i>quant, rest_ind, psoa, index</i>
<i>quant</i>	SEMDET = <i>atom</i> REST_IND = <i>rest_ind</i>	
<i>rest_ind</i>	INDEX = <i>index</i> RESTR = <i>restr</i>	<i>noun_obj, pred_obj</i>
<i>index</i>		<i>event_index, ind_index, location_index, expl_index</i>
<i>ind_index</i>	PERS = <i>atom</i> GEN = <i>atom</i> NUM = <i>atom</i>	<i>ev_ind_index</i>
<i>event_index</i>	AKT = <i>atom(stat,non_stat)</i>	<i>ev_ind_index</i>
<i>ev_ind_index</i>		
<i>location_index</i>		
<i>expl_index</i>		
<i>restr</i>	PAS = <i>psoa</i> SEM_ADJ = <i>set of contents</i>	
<i>psoa</i>	REL = <i>atom</i>	<i>arg0, a1, a2, aADJ</i>

There basically are two aspects that are worth commenting in relation to this type system: the treatment of the *index* type and the definition of the value of the feature REL.

Since we are proposing an extension to the standard use of indexes in HPSG, there appears much more complexity in the definition of this type than it used to be before. First of all, we foresee at least four distinct types of index: the individual index (which is the standardly used one), the index for eventualities, another one for locations (which is needed to represent in an appropriate way the locative arguments of predicates (see, for example, in section 4.2.3.7 an example with such an index), and the expletive index).

The question now arises as to what structure is given to the new eventive index. First of all, the index seems to be the appropriate place to indicate the *Aktionsart* value (namely, stative or not) of the construction; recall that this property at the phrasal levels is related to the quantity properties of the arguments the predicate has and is thus related to the referential properties of the elements in the construction. And secondly, the eventive index is shared by both verbal predicates and predicative nouns; the latter however also share some properties with the rest of nouns (namely, person and number), which are characteristically stated in individual indexes. We are thus led to a more complex type structure, in which both the *ev\_index* type and the *ind\_index* type have a common subtype (*ev\_ind\_index*), which is appropriate for predicative nouns. This is a limited multiple inheritance which helps explaining the properties of the indexes of predicative nouns (which have some characteristics of nominals and some of predicates).

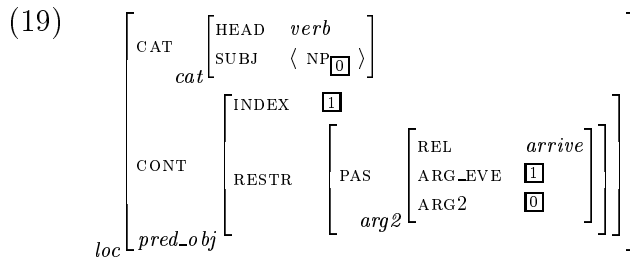
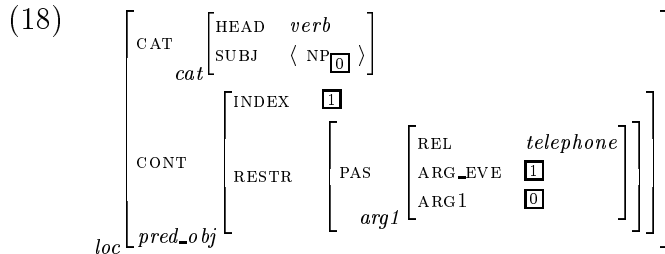
As characterised in the type hierarchy above, the feature REL has an atomic value. This is an option which has been taken in order to maintain the simplicity of the type system; were we to characterise relations as types (as it is suggested in Pollard & Sag, 1994) we would end up with a fairly complex inheritance network. On the other hand, it is worth mentioning that a still more complex structure can be given to that value, if some additional characterisation is wanted. This has been exemplified in this paper with the labelling of adjuncts (see section 8.4 below), were a complex type is introduced as the value of the feature REL in order to state the class of modification the relation is introducing. Such an approach could also be followed with the other types of relations if we were interested in classifying them according to some semantic system.

## 4 Exemplification: The Lexical Representation of Verbs

### 4.1 Introduction

In TFS formalisms there is no indeterminacy produced in parsing as to the mapping from syntactic relations to the ones determined in the content part of the sign. Indeed, there is a unique lexical specification for each verb (or verb reading), in which the value of the content of the syntactically specified complement is unified with the value of the corresponding argument slot in the PAS. This can be easily seen in the following two schematic lexical entries (for *telephone*, an unergative verb, and for *arrive*, an unaccusative verb), which together exemplify the one-to-many mapping from the subject (i.e., the unique

member of the SUBJ list) to the first or the second argument (labelled either ARG1 or ARG2 in the *psoa* value).<sup>9</sup>



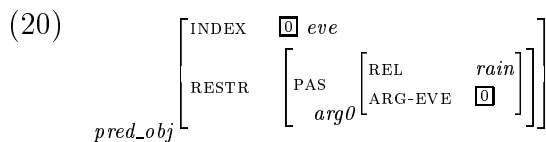
As can be easily seen it is the lexical entry of the verb that establishes the link between the syntactic constituent and the argument slot it occupies in the PAS. This obviously will not create any kind of indeterminacy in the analysis process, as the mapping is uniquely determined in the lexical entry.<sup>10</sup>

Following the ideas on the type system developed in section 3 above and taking into account the previous remark on the mapping from syntactic structure to PAS, we present in the next sections a proposal for the lexical representation of the different verb classes.

## 4.2 Phrasal Arguments to Verbs

### 4.2.1 0-Valent Verbs

The plain, direct use of 0-valent verbs is the one in which, except for the ARG-EVE no argument is present. This means that in the PAS included in the verbal sign there is a relation with only the ARG-EVE, as shown in the following:



<sup>9</sup>Note that these are the *local* values of general lexical entries (they do not specify the properties that stem from the morphological form of the predicate –thus there is no information about tense and aspect, nor about the properties of the subject, if any); henceforth we follow this practice with respect to all lexical entries.

<sup>10</sup>Note however that the indeterminacy remains that is induced by the presence of more than one applicable lexical entry.

As this implies that the syntactic side of the sign indicates that there is no complement subcategorised for by these verbs their subject list is going to be empty in *pro-drop* languages and filled with an expletive subject in the rest. Thus a schematic entry for **es: llover** is going to be as the following:

$$(21) \left[ \begin{array}{l} \text{CAT} \\ \text{cat} \\ \text{CONT} \\ \text{pred\_obj} \\ \text{loc} \end{array} \left[ \begin{array}{l} \text{HEAD } verb \\ \text{SUBJ } \langle \ \ \rangle, \text{COMPS } \langle \ \ \rangle \\ \text{INDEX } \boxed{0} \text{ eve} \\ \text{RESTR } \left[ \begin{array}{l} \text{PAS} \\ \text{arg0} \end{array} \left[ \begin{array}{l} \text{REL} \\ \text{ARG-EVE} \end{array} \left[ \begin{array}{l} \text{lllover} \\ \boxed{0} \end{array} \right] \right] \right] \end{array} \right] \right]$$

Whereas *rain* would be represented as follows:

$$(22) \left[ \begin{array}{l} \text{CAT} \\ \text{cat} \\ \text{CONT} \\ \text{pred\_obj} \\ \text{loc} \end{array} \left[ \begin{array}{l} \text{HEAD } verb \\ \text{SUBJ } \langle \left[ \begin{array}{l} \text{CAT } | \text{ HEAD } \textit{noun} \\ \text{CONTENT } | \text{ INDEX } \textit{expl\_index} \end{array} \right] \rangle, \text{COMPS } \langle \ \ \rangle \\ \text{INDEX } \boxed{0} \text{ eve} \\ \text{RESTR } \left[ \begin{array}{l} \text{PAS} \\ \text{arg0} \end{array} \left[ \begin{array}{l} \text{REL} \\ \text{ARG-EVE} \end{array} \left[ \begin{array}{l} \textit{rain} \\ \boxed{0} \end{array} \right] \right] \right] \end{array} \right] \right]$$

It is interesting to note that in the lexical entry for **es: llover** the value of the SUBJ attribute is the empty list ( $\langle \ \ \rangle$  in the feature structure above); and that in the lexical entry for *rain*, it is a list with a single element which is an expletive index.

In some languages these verbs allow for an argument, as exemplified in the following:

- (23) a. **de:** Es regnet Blüten  
**es:** Llueven reproches  
 (There are masses of counterfeits)
- b. **de:** Es hagelt Vorwürfe  
 (The accusations come thick and heavy)

In these cases, however, a different meaning is given to the lexical item and therefore a new lexical entry is needed to code its PAS. Note that in *pro-drop* languages these predicates behave like unaccusative intransitive verbs, when an argument is introduced. Note also that this kind of verbs do not allow for referential arguments. As a consequence in the PAS typing we are using they would be typed as *arg2*.

#### 4.2.2 Monovalent Verbs

Two basic classes of intransitive are distinguished:

- unergatives



- (24) a. John telephoned  
       **it:** Giovanni telefona  
       b. John played  
       c. John walked

- unaccusatives

- (25) a. John arrived  
       **it:** Giovanni arriva

We assume the unaccusative hypothesis, that is, that not all monovalent verbs have a single type of argument: unergatives have an ARG1 as their sole argument and unaccusatives have just an ARG2 as the surface subject is taken to be the deep object of the verb at some underlying level.

It is true that the unaccusative hypothesis presents problems, particularly the lack of strict criteria for deciding which verbs are unaccusatives and which are unergatives for all languages. Although it has always been claimed that the unaccusative hypothesis is syntactic; at the same time it has usually been felt that there is a semantic basis for the initial grammatical relations (expressed within Relational Grammar as the “Universal Alignment Hypothesis”). However it has now repeatedly been pointed out that criteria do not yield always a unique result, even for languages having clear, established formal criteria for distinguishing the two types of intransitives (like Dutch and Italian) (Zaenen, 1988; Sanfilippo, 1990). In Sanfilippo’s account it is true that there is a semantic basis for the distinction; nonetheless it is not a unique semantic factor (namely the thematic structure of the predicate) that explains the syntactic distinction, but there are at least two such semantic factors (the thematic and the aspectual structures of the predicates).

However this characterisation of the first argument as ARG1 or ARG2, particularly when it is the unique argument, is closer to a fully semantic typing and tends to help in the multilingual mapping between lexical (and phrasal) representations. Note also that under this approach the relation between ergative verbs and their unaccusative version (whether it entails phonological, morphological or syntactic changes, or not) can be easily represented. In a monotonic system this kind of relation can only be expressed if the basic information is common to the two types of signs. The only change between the two classes of lexical entries is the introduction (or elimination) of the external argument (which is typed as ARG1), with the other, internal arguments remaining as they are (see 4.2.3.1 below). The basic representations for intransitive, unaccusative and unergative verbs can be seen in the previous subsection 4.1.

### 4.2.3 Divalent Verbs

**4.2.3.1 Subject-Object — Dynamic** The prototypical verbs included under this heading are classic, passivisable, transitive verbs (like *eat*, *hit* or *mark*). In the following

the standard representation for them is exemplified (for the non morphologically marked forms of *mark*):

$$(26) \left[ \begin{array}{l} \text{CAT} \\ \text{cat} \\ \text{CONT} \\ \text{loc } \text{pred\_obj} \end{array} \left[ \begin{array}{l} \text{HEAD } verb \\ \text{SUBJ } \langle \text{NP}[\textit{nom}]_{\boxed{0}} \rangle, \text{ COMPS } \langle \text{NP}[\textit{acc}]_{\boxed{1}} \rangle \\ \text{INDEX } \boxed{2} \textit{eve} \\ \text{RESTR } | \text{ PAS } \left[ \begin{array}{l} \text{REL } \textit{mark} \\ \text{ARG-EVE } \boxed{2} \\ \text{ARG1 } \boxed{0} \textit{content} \\ \text{ARG2 } \boxed{1} \textit{content} \end{array} \right] \\ \text{arg1\_2} \end{array} \right] \right]$$

Although there are clear differences in  $\theta$ -role assignment among the verbs in this class we stick to current linguistic practice and consider the first and second argument as semantically unrestricted: therefore we allow ARG1s which are agents, causers or experiencers; and ARG2s which are patients, themes or experiencers. This treatment of dynamic divalent verbs implies that accusative, dative and prepositional complements are treated alike. See the following examples for a verb subcategorising for a dative complement (**de**: *helpen*) and for one that takes a prepositional object (**fr**: *se souvenir de*):

$$(27) \left[ \begin{array}{l} \text{CAT} \\ \text{cat} \\ \text{CONT} \\ \text{loc } \text{pred\_obj} \end{array} \left[ \begin{array}{l} \text{HEAD } verb \\ \text{SUBJ } \langle \text{NP}[\textit{nom}]_{\boxed{1}} \rangle, \text{ COMPS } \langle \text{NP}[\textit{dat}]_{\boxed{2}} \rangle \\ \text{INDEX } \boxed{0} \textit{eve} \\ \text{RESTR } | \text{ PAS } \left[ \begin{array}{l} \text{REL } \textit{helpen} \\ \text{ARG-EVE } \boxed{0} \\ \text{ARG1 } \boxed{1} \textit{content} \\ \text{ARG2 } \boxed{2} \textit{content} \end{array} \right] \\ \text{arg1\_2} \end{array} \right] \right]$$

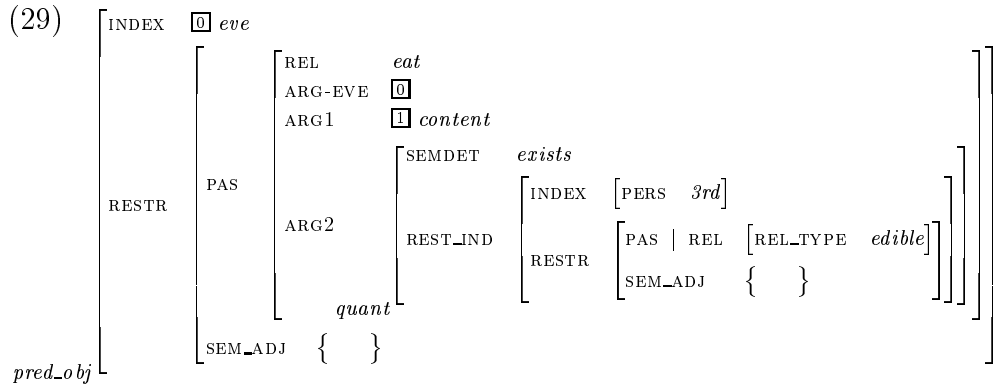
$$(28) \left[ \begin{array}{l} \text{CAT} \\ \text{cat} \\ \text{CONT} \\ \text{loc } \text{pred\_obj} \end{array} \left[ \begin{array}{l} \text{HEAD } verb \\ \text{SUBJ } \langle \text{NP}[\textit{nom}]_{\boxed{1}} \rangle, \text{ COMPS } \langle \text{NP}[\textit{obliq}; \textit{de}]_{\boxed{2}} \rangle \\ \text{INDEX } \boxed{0} \textit{eve} \\ \text{RESTR } | \text{ PAS } \left[ \begin{array}{l} \text{REL } \textit{se souvenir} \\ \text{ARG-EVE } \boxed{0} \\ \text{ARG1 } \boxed{1} \textit{content} \\ \text{ARG2 } \boxed{2} \textit{content} \end{array} \right] \\ \text{arg1\_2} \end{array} \right] \right]$$

Many of these verbs allow for alterations of their subcategorisation lists. The major alterations are due to the elimination of one of their arguments. There are basically three distinct cases to be considered:

- object deletion verbs

Verbs like *eat*, *drink* and *telephone* can be realised with or without the second argument. A lexical entry with an empty COMPS list (i.e., with only the subject) can account for such occurrences. If we are interested in distinguishing these intransitive readings from true intransitive verbs, in the content part of the sign there can

be specified that the second argument of the predicate is unexpressed but implied, by the introduction of a non restricted existential quantification. See for example how the *content* of the intransitive reading of *eat* could be represented:



- ergative verbs

Verbs like **it:** *affondare* and **es:** *mejorar* can be realised as well as transitive and as intransitive. In this case however the  $\theta$ -role of the unique complement in the intransitive realisation is identical to the one of the second argument in the transitive pattern. A lexical entry for the intransitive realisation will include only one complement in the SUBJ list (i.e., the subject) and its content will be linked to the value of the ARG2 attribute in the content part of the sign. Thus such an intransitive realisation of these verbs will be practically identical to that of the “true” intransitives that are unaccusatives (see section 4.2.2 above). Consequently, the two lexical entries are linked with one another in a different way as the ones of object deletion verbs; here the external argument is removed and as a consequence the internal argument (ARG2 in our typing) occupies the subject position in the subcategorisation lists.

Note that in the examples just mentioned the lexical relation is established between two kinds of lexical entries which differ in the value of the COMPS feature and in the structure of the PAS in the content part of the sign, but that there is not change in phonological (or orthographic) form implied:

- (30) a. **es:** El consejo mejora la situación  
(The council improves the situation)
- b. **es:** La situación mejora  
(The situation improves)
- (31) a. **it:** L’artiglieria ha affondata due navi  
(The artillery has sunk two ships)
- b. **it:** Due navi sono affondate  
(Two ships have sunk)

In the last example shown there is a difference in the auxiliary selection of the verb **it:** *affondare*, according to its PAS realisation. This however need not be stated

in the particular lexical rule relating the two entries for this verb, since it can be derived from the general rules of auxiliary selection in Italian (which roughly say that unaccusatives select *essere*).

There are however verbs which relate in exactly the same way (both syntactically and semantically) and nonetheless do not have the same surface form:

- (32) a. **fr:** Le conseil améliore la situation  
(The council improves the situation)  
b. **fr:** La situation s'améliore  
(The situation improves)
- (33) a. **de:** Der Feind hat einen großen Teil der Flotte versenkt  
(The enemy has sunk an important part of the fleet)  
b. **de:** Ein großer Teil der Flotte ist versunken  
(An important part of the fleet has sunk)

These two examples show that there are two different sorts of changes in surface form: it is of a grammatical nature (as in the French example, in which the difference in the PAS value of the verb is marked with the clitic *se*), or it is lexical (as in the German example, where the difference is marked with a change in the tonic vowel of the lexical root). Of course the possibility of fully dealing with the latter cases depends on the power of the morphological analyser and on the relation between morphological operations and those structuring the lexicon. However the fact remains that the syntactic and semantic relations are identical.

- alterations implying the absence or de-thematisation of the external argument

The basic operations are “passivisation” and “impersonalisation”. In passivisation the external argument is de-thematised, so that it only appears as an explicit PP complement headed by *by*. On the other hand, impersonalisation implies that the external argument is not present (either through an agentless passive or through an explicitly impersonal construction). In this latter case it would be desirable to have a mechanism to refer to the ARG1; we suggest a treatment similar to the one proposed by object deletion verbs.

In passive constructions the SUBJ list and the COMPS list of the verbal predicate would contain respectively one NP complement: the first one, the subject, would have its semantic value coindexed with the value of the ARG2 feature of its PAS; and the second one, marked with preposition *by* would be coindexed with the value of the ARG1 feature.

On the other hand, agentless passives could be treated in a similar way to object deletion verbs: the subject would be coindexed to the ARG2 value of the verbal PAS, and the ARG1 feature would be filled with a non restricted index (with an existential quantifier). In this way agentless passives could obtain a PAS representation similar to the one of impersonals with indetermined subjects (e.g., with *people*, **de:** *man*, **fr:** *on*, and so on).

**4.2.3.2 Subject-Object — Stative** This constitutes a subclass of transitive verbs which is characterised by the fact that they do not allow passive construction.

Under the current framework they receive an identical representation to that of dynamic predicates; thus the lexical entry for *resemble* would be like the following:

$$(34) \left[ \begin{array}{l} \text{CAT} \\ \text{CONT} \\ \text{loc} \end{array} \left[ \begin{array}{l} \text{HEAD } verb \\ \text{SUBJ } \langle \text{NP}[\textit{nom}]_{\boxed{0}} \rangle, \text{ COMPS } \langle \text{NP}[\textit{acc}]_{\boxed{1}} \rangle \\ \text{INDEX } \boxed{3} [\textit{AKT } stat] \\ \text{RESTR } | \text{ PAS } \left[ \begin{array}{l} \text{REL } resemble \\ \text{ARG-EVE } \boxed{3} \\ \text{ARG1 } \boxed{0} content \\ \text{ARG2 } \boxed{1} content \end{array} \right] \end{array} \right] \right]$$

The fact that dynamic, and therefore passivisable, verbs are represented in the same way as stative, non-passivisable predicates creates a difficulty in the formulation of the lexical rules stating the active-passive relationship. This can be avoided in the current framework if we distinguish between two classes of eventualities by means of a feature AKT (for *Aktionsart*) which has two possible values: *stat* for the stative predications, and *non\_stat* type for the non stative ones.

As a consequence the active-passive relationship would be available for verbs complying with conditions stated in the syntax and in the content part of the sign. Thus the passivisation could occur only if the lexical sign is verbal and subcategorises for an accusative complement (both conditions are stated in the syntactic part of the sign), and if it introduces an index over non states (which is stated in the content part).

**4.2.3.3 Attribute of Subject** In principle there are two possibilities to treat attributes to the subject: either as the unique complement of a raising verb, or as the second, predicative argument of an *equi* one. In the second case the main predicate is interpreted as a control verb, so that the external argument (i.e., the ARG1) of the attribute is structure-shared with the content of the subject of the main predicate (which would have a PAS with two arguments: ARG1 and ARG\_AS). This treatment however makes only sense in the present framework if there are verbs which both assign a  $\theta$ -role to their subjects and still have an attribute of the subject. This of course does not apply to the copula, which clearly does not assign  $\theta$ -role to its subject. Nonetheless, we maintain in the type system (see section 3) the type *arg1\_AS* for the cases (if any) in which an attribute to the subject appears with a verb that assigns a  $\theta$ -role to its subject.

In what follows we discuss the representation of attributes of the subject in relation to the copula. Since the copula does not assign  $\theta$ -role to its subject, we assign to all its attributes an identical representation and a unique label: ARG2. This treatment applies to stative predications of the subject (whatever is their syntactic category) and also to locative predications:<sup>11</sup>

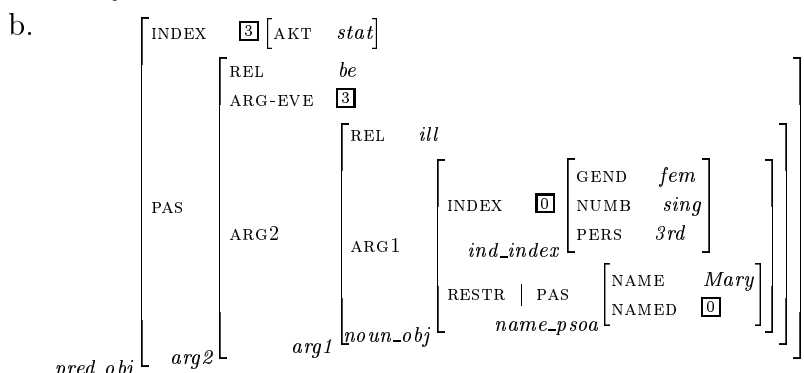
<sup>11</sup>But it does not apply to equative constructions, which are treated as stative transitives.

- (35) a. Mary is / feels ill  
 b. Mary is in pain  
 c. Mary is in a bad mood  
 d. Mary is a doctor
- (36) Mary is in the city

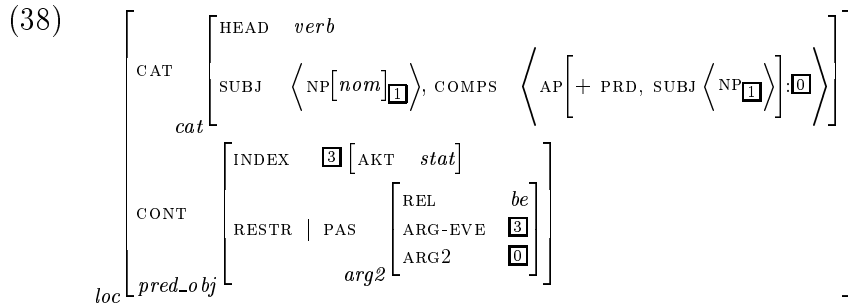
There is in the type system developed a clearcut distinction between predicative and non predicative contents. The latter class is restricted to referential and quantified NPs, thus leaving all other semantically full expressions as predicative. Predicative means that the sign can occur in a predicative position, i.e as a subject or object attribute. Most of the constituents that can occur in a predicative position can also occur as adjunct to nominal or to verbal signs. This is the case of many adjectives, so we propose to distinguish between different sorts of adjectives and to assign them, when they can appear in predicative position, an external argument (see section 5 on adjectives). A similar distinction between different sorts of PPs and NPs seems appropriate. We assume all PPs and NPs that occur with the copula (in other words, as attribute of subject) or as attribute of object (see subsection 4.2.4.3 on object attributes) to be predicative and to have therefore assigned an external argument. In order to account for the two possible positions (adjunct and attribute position) of these constituents we consider the possibility of establishing links between the two lexical entries via lexical rules (see 8.2.1 on adjuncts to nominal signs).

Let us first concentrate on predicative adjectives (i.e., those that can occur in post-copular position). The verb *be* does not assign thematic role to its subject, so that the only way for that complement to obtain a  $\theta$ -role is that it be assigned by the only remaining predicative element, that is to say, by the adjective. As a consequence it seems adequate to treat *be*, and its aspectual variants (*become*, *remain*, and so on), as raising verbs, in a way that is consistent with the treatment proposed for *seem*. Thus in a sentence like the one in (37a) the final PAS would be as shown in (37b).

- (37) a. Mary is ill



In order to achieve this the adjective *ill* would be lexically represented as indicated in section 5 on adjectives, and the lexical entry for *be* would treat it as a raising to the subject verb:



Note that this representation is absolutely similar to that of other raising to the subject verbs (like *seem* — see section 4.3.1.4). Note also that there must be a lexical relation between the predicative use of adjectives and their appearance as NP internal constituents, since although it is not the general rule that adjectives can occur in both positions there are many of them that can (see section 8.2.1).

The other two categories that usually occur in attribute to the subject position are PPs and NPs. What is needed to generalise to them the treatment proposed is that the NPs and PPs that can occur in this positions are labelled PRD + and have a one-member list as value to the SUBJ attribute. For example, bare plurals of countable nouns and bare singulars of mass nouns in English can appear in this position, and therefore there should be a lexical entry for them marked as predicative (of course this entry might be the result of the application of a lexical rule). Indeed the conditions have to worked out for each language for NPs and PPs to occur in this position.

**4.2.3.4 Prepositional Verbs** As has already been mentioned in section 4.2.3.2 on Subject-Object verbs, prepositional verbs are treated exactly in the same way as subject-object verbs are. Thus a verb like *rely on* will be given a PAS with two arguments, labelled ARG1 and ARG2, respectively. For an example of a lexical entry see the one for **fr**: *souvenir de* above (section 4.2.3.2).

Note first that the preposition is a strongly bound one, and that therefore it is not treated as the head of the constituent, but simply as a marker to it; and secondly, that the preposition cannot be seen as independent of the NP it marks (that is to say, it tends to form a constituent with the NP, and not with the verb).

It must be taken into account as well that many of these verbs can take sentential complements. In some languages the preposition still marks the complement, whereas in other ones it simply disappears.

- (39) a. **es**: Juan confía en María  
           (Juan relies on Mary)  
       b. **es**: Juan confía en que no lloverá  
           (Juan relies on the fact that it won't rain)
- (40) a. Mary agreed that they should meet

- b. **fr:** Il se souvient de toi  
(He remembers you)
- c. **fr:** Il se souvient que tu le lui as dit  
(He remembers that you have told him)

Finally a note on passivisation. In languages like English that allow for preposition stranding, these verbs can form “pseudo-passives”:

- (41) a. This matter was looked into (by the Commission)
- b. This position was argued for by the Commission
- c. John was counted on by his colleagues
- d. The money was asked for by Mary
- (42) That they should meet was agreed on

These passive constructions can be accounted for in active-passive lexical relations, in a similar way to the one mentioned above; it has only to be taken into account that for this kind of languages the syntactic constraints to passivisation have to be loosened with respect to the general rule stated above.

With respect to this group of verbs, care has to be taken to distinguish them from phrasal verbs, in which the preposition tends to form a constituent with the verb and not with the NP, and from verbs taking ARG\_ADJs, where the preposition is not a strongly governed one (see the three following subsections).

**4.2.3.5 Phrasal Verbs** Three types of phrasal verbs can be distinguished according to the kind of preposition or particle that they take:

- grammatical preposition

(43) The system looks up the words in a database

- semantic preposition

(44) He switched the light off

- completive preposition

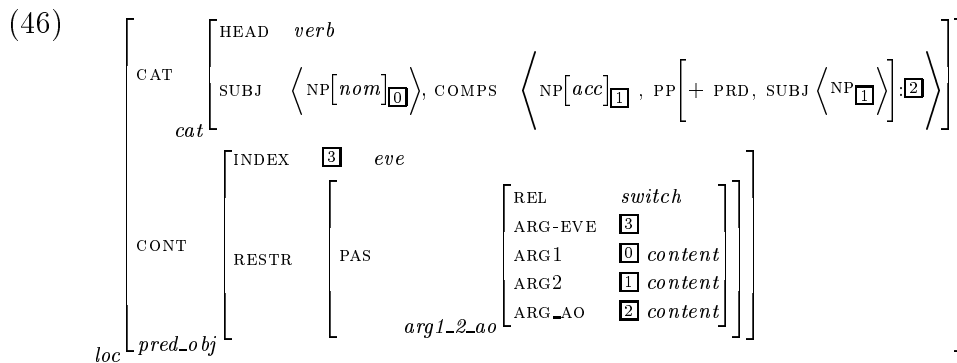
(45) He tore the paper up

There are two possibilities to treat phrasal verbs. The first one is to treat them as lexicalised, so that for the examples above we would have the following lexical entries: *look\_up*, *switch\_off* and *tear\_up*. In our framework this would imply to treat the verb and the particle as a constituent, which may imply the recognition of discontinuous constituents. Were this approach chosen, phrasal verbs would be treated as any other



two argument verb. However this seems not to be a suitable analysis at least for the class of phrasal verbs with a semantic preposition but it could be advocated on the ground of translational simplicity.

The other possibility, which seems a more correct analysis for the class with a semantic preposition, is to treat phrasal verbs as three argument verbs by considering the particle a predicative third argument. In this case the particle would have –as a predicative argument– an external argument which would be coindexed with the second argument of the main predicate. See for example the simplified entry for *switch* in a sentence like *He switched the light off*:



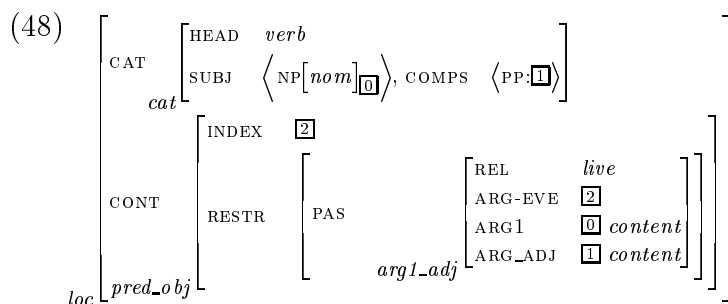
Linguists therefore have to decide how to implement this group of verbs: as lexicalised phrasal ones (in which case they are treated as any other two arguments verbs at the level of PAS, and the syntactic operation of recognising the particle has to be defined), or as three argument verbs with a predicative third argument (which has to be treated as other predicative third arguments —see section 4.2.4.3).

**4.2.3.6 Arg-Adjunct (Stative)** The standard sort of predicates falling under this heading are stative verbs taking a PP with a weakly bound preposition, usually with a locative meaning:

- (47) a. She lives in Budapest  
 b. In fact, she lives just inside the city boundaries  
 c. **de:** Das Buch liegt unter dem Bett  
 d. **fr:** La mairie se trouve au centre ville

As the first two sentences show, the preposition is variable and modifiable, and semantically contentful. Therefore it has to be represented in the PAS in the content part of the phrasal sign. The lexical verbal sign has to take into account this fact, and this in practice means that the predicate subcategorises for a PP (and not merely for a prepositionally marked NP, as in the case of prepositional complements), and that the second argument in the PAS is asked to be a relation of the locative sort. See for example the simplified entry for *live*:<sup>12</sup>

<sup>12</sup>See the discussion of weakly bound prepositions in section 2.4.



In the current framework we introduce an ARG\_MEASURE label. This sort of complement is an exceptional case both with respect to the potential syntactic realization (NP or PP) and with respect to the applicability of the criterion which distinguishes normally ARG\_ADJUNCTs from ARG2s (weakly or strongly bound preposition). This means that even NPs or PPs with strongly bound prepositions, no matter which valency the verb has, can be assigned ARG\_MEASURE. See the following examples with a NP realization and with a strongly bound preposition respectively:

- (49) a. The Dutchman measured 2 metres  
 b. The meeting lasted for two hours

There is a clear semantic similarity between both cases, which is best captured by a unique argument label. Note that in this case the PP is to be treated as a NP and not as a predicative PP.

We also foresee the use of the label ARG\_MANNER for some argument-structure alternations of transitive verbs. In particular, there is a set of verbs which can undergo the process sometimes called of decausativisation which results in a middle construction:

- (50) a. They sold very few books last night  
 b. This book sells well  
 c. \*This book sells

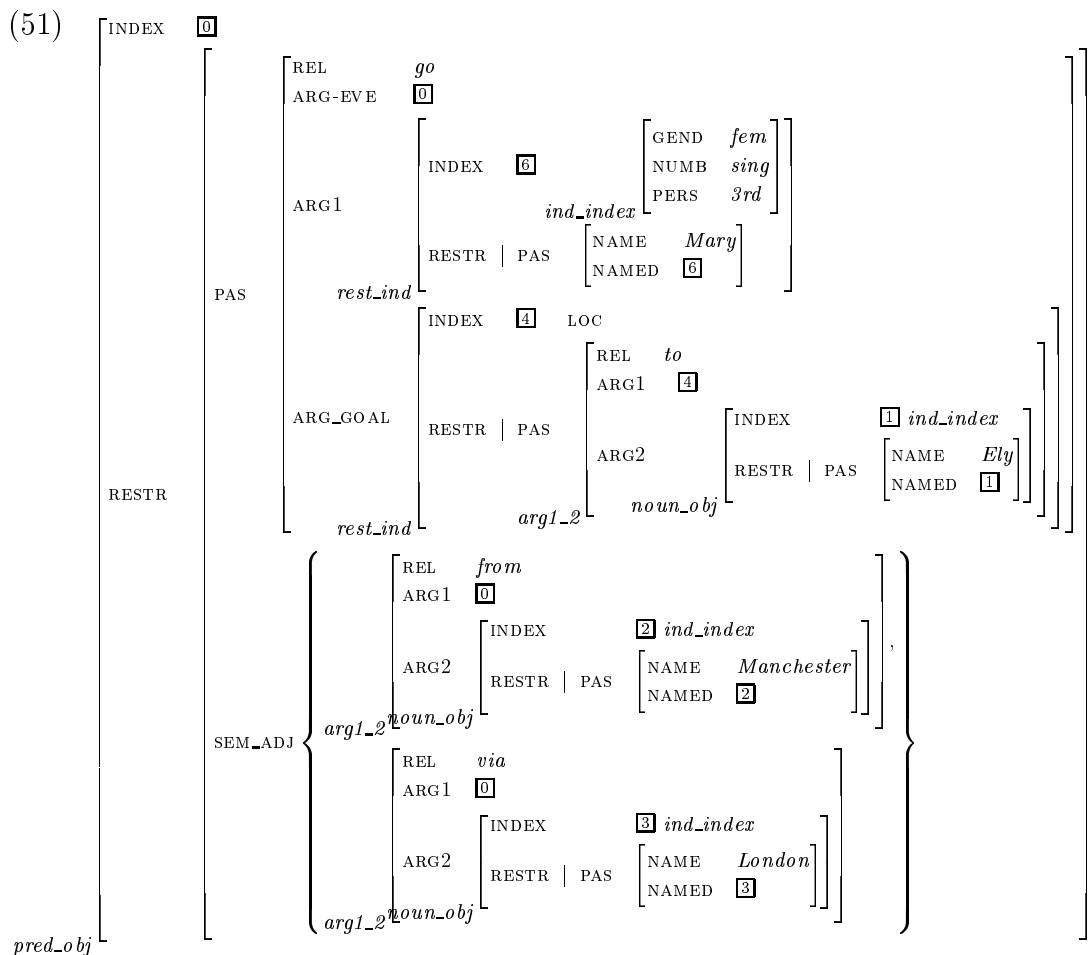
The verb *sell* has an ordinary transitive reading, exemplified in (50a). One of the possible alternations that affect this verb is decausativisation (50b); however if this process is undergone a manner adverbial has to be present (50c), which therefore has to be considered as subcategorised for in this reading.

Although we have not been able to find an example of such a complement in the primary reading of a verb, the fact remains that there are alternations to transitive verbs in which manner adverbials like this are obligatory. We therefore foresee such a label in our set of ARG\_ADJUNCTs.

**4.2.3.7 Arg-Adjunct (Dynamic)** With respect to dynamic predicates there is no added difficulty to maintaining the same approach on the lines followed for stative predicates. Although the status of directionals within argument theory has always caused

difficulties, the position adopted in the present framework is the following: all directional complements of basic intransitive change-of-location verbs will be considered as having an ARG\_ADJUNCT label (namely ARG\_ORIGIN or ARG\_GOAL). E.g. in the case of *to go* the element expressing the goal of the activity is most likely to appear with the verb, in other cases the origin (or may be both) might be focussed by the semantic class the verb belongs to. Note that in the proposed treatment locative PPs (and ADVPs) are represented similarly when they are taken to be ARG\_ADJs, predicative complements or modifiers (except for their MOD attribute — see section 8 on adjuncts). Their place in PAS however is radically distinct, since the two former appear as attributes in the PSOA type (i.e., the value of the attribute PAS) and the latter only appear in the set which is the value of the SEM\_ADJ attribute.

On the basis of what we said above in a sentence like *Mary went to Ely from Manchester via London*, *Mary* is assigned ARG1; *to Ely*, ARG\_GOAL; *from Manchester*, MOD; and *via London*, MOD as well.<sup>13</sup> This sentence would be given the following (partially specified) PAS in the current framework:



<sup>13</sup>Note that these two modifiers could be labelled origin and path respectively, if the proposal in section 8.4 below is adopted.

**4.2.3.8 Resultative Secondary Predications** The secondary predications in the current framework can be treated without any problem as such (with the secondary predication marked as PRD +, see section 2.4). In general two sorts of such structures can be distinguished: *equi* and raising structures. Resultative secondary predications can be exemplified by the following:

(52) The ice-cream froze solid

In this example the verb *freeze* selects the subject *the ice-cream* and the secondary predicate *solid*; and this latter predicate selects as well the subject *the ice-cream*. Therefore an *equi* PAS structure seems to be the one that best captures these facts, like in the following simplified lexical entry for *freeze* with a secondary resultative predication:

(53)

CAT	[	HEAD	verb	,	SUBJ	⟨ NP[nom] <sub>[0]</sub> ⟩,	COMPS	⟨ AP [ + PRD, SUBJ ⟨ NP <sub>[0]</sub> ⟩ ] : <sub>[1]</sub> ⟩	]											
CONT	[	INDEX	[ 2 ] eve	,	RESTR	[	PAS	[	REL	freeze	,	ARG-EVE	[ 2 ]	,	ARG2	[ 0 ] content	,	ARG_AO	[ 1 ] content	]
loc	pred_obj																			

**4.2.3.9 Unaccusative + Arg-adjunct** Finally, let us note that there are some ergative verbs that allow for an ARG\_ADJ. In the case that they have their full argument structure represented they will have an ARG1 realised as the subject, an ARG2 realised as an object, and an ARG\_ADJ:

(54) Mary poured the water from the vase

However, if they are realised as unaccusative predicates they will have only two arguments: the subject (which is an ARG2) and the ARG\_ADJ:

(55) The water poured from the broken vase

This means that we have to allow in the PAS type structure for verbs having an ARG2 and an ARG\_ADJ, like the following:

(56)

CAT	[	HEAD	verb	,	SUBJ	⟨ NP[nom] <sub>[0]</sub> ⟩,	COMPS	⟨ PP: <sub>[1]</sub> ⟩	]											
CONT	[	INDEX	[ 2 ]	,	RESTR	[	PAS	[	REL	pour	,	ARG-EVE	[ 2 ]	,	ARG2	[ 0 ] content	,	ARG_ADJ	[ 1 ] content	]
loc	pred_obj																			

#### 4.2.4 Trivalent Verbs

**4.2.4.1 Object + Arg-Adjunct** The condition that only PPs headed by weakly bound prepositions can be labelled as ARG\_ADJs has to be relaxed for some tri-valent (and tetra-valent verbs). Consider the following examples:

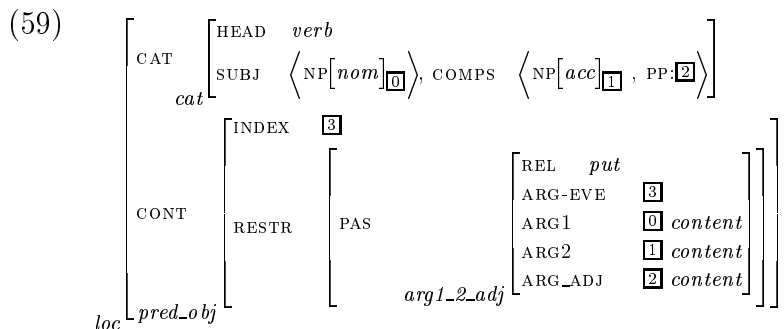
- (57) a. John emptied petrol from the tank  
 b. **de:** Der Richter wendet das Gesetz auf den vorliegenden Fall an  
 c. John split the group into three teams

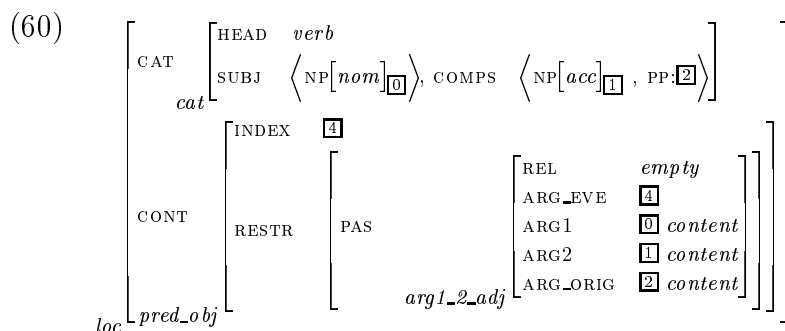
Applying the above mentioned criterion, the PP should become ARG2 as it is headed by a strongly bound preposition. This however, is not possible, as the direct object is already labelled ARG2. Thus in these cases one has to check first whether the label ARG2 can or should be assigned to one of the arguments. If it is so the PP has to receive an ARG\_ADJ label, although the preposition is strongly bound. Note that an ARG1-ARG2E-ARG2 analysis is neither appropriate for these examples, because such an analysis is used only for thematic ARG2 and non-thematic ARG\_2E as in the following example:

- (58) John emptied the tank of petrol

In addition such an analysis would raise problems in connection with the unaccusative hypothesis, as the surface subject in the anti-causative alternant of sentence (57c) would be an ARG\_2E, and not an ARG2.

A second question is whether the ARG\_ADJ is also in this case to be treated as a secondary predication. According to what we said above (see subsection 4.2.3.3 on attributes of the subject) we propose to treat all locative PPs in predicative position (subject or object attributes) as having an external argument, just as PRD + elements are, but not locative PPs labelled as ARG\_ADJ. Just for the purpose of exemplification, we propose the (simplified) lexical entries for *put* and *empty*.





Most of these verbs allow for alterations of their PAS, basically through the passivisation and impersonalisation processes. The relations between active and passive realisations can be formulated in exactly the same way that they were with simple subject-object verbs.

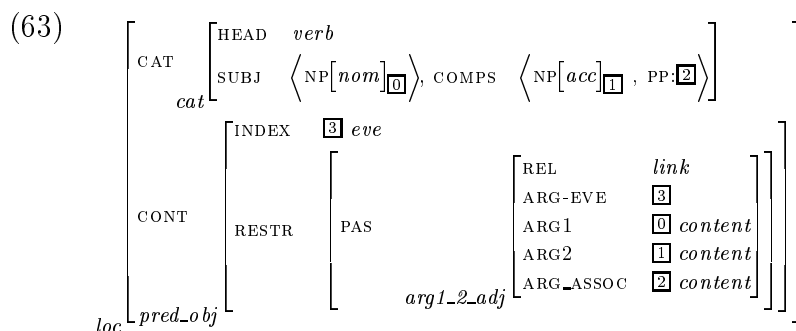
**4.2.4.2 Object + Arg-Adjunct; relational** A particular type of ARG\_ADJ (namely ARG\_ASSOC) is introduced in order to account for the complements of verbs like *link* or *compare*. These verbs allow for a double realisation of their non external arguments: either with two complements or with a unique complement in plural. In the former case, one complement is thematic and is realised by an NP in accusative case, and the other one is a PP introduced by *with* (in some languages it can also be introduced by the preposition *to*, or its equivalents). This realisation is exemplified by the following sentences:

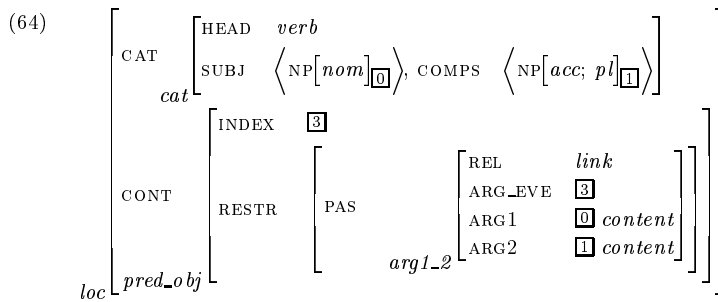
- (61) a. The Suez Canal links the Mediterranean Sea with/to the Persian Gulf  
 b. He compared Peter with/to Hans

These verbs however can have a unique internal complement, which has to be in plural:

- (62) a. The Suez Canal links the Mediterranean Sea and the Persian Gulf  
 b. He compared the two boys

Here we present a couple of lexical entries for *link*, which exemplify the two uses just mentioned.





**4.2.4.3 Object attributes** Object attributes are characterised as PRD + elements, just as subject attributes are; in fact the same sort of elements can occur in both positions (which may be AP, NP, PP or non finite constructions).<sup>14</sup> Note that attributes to objects, like attributes to subjects, can be both optional and obligatory. In the following sentence, for example, the object attribute is an obligatory AP complement to the verb:

(65) Mary considered herself happy

There is however a related class of elements which are omissible secondary predications and should be treated as modifiers, since they are not subcategorised for by the verb.<sup>15</sup>

See the following examples:

- (66) a. Mary served the meal hot  
 b. John ate the meat raw

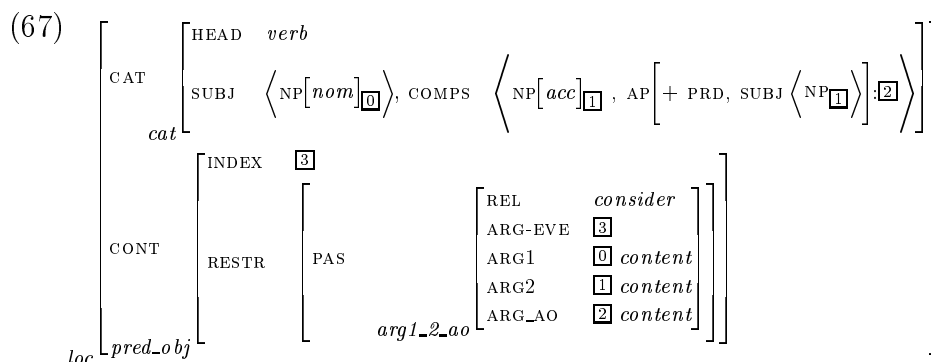
It only remains to be stated that for all sorts of constituents the verbal predicate has to be taken as an *equi* one (just as *freeze* was treated in the subsection on resultative secondary predications).

Note that the problems that exist in characterising the predicative elements that can occur in this position are exactly the same we mentioned in trying to determine the class of elements that can appear as attribute to the subject.

Just for exemplification we present in the following the simplified lexical representation of *consider*:

<sup>14</sup>Consider as well the section 4.3.1.5 on small clauses.

<sup>15</sup>Of course, if we do not want to treat these elements as modifiers we should have two different lexical entries for many verbs, i.e. for the *EQUI* reading and for the transitive one. Note also that there are some modifiers which are semantically ambiguous, so that they can have an object-oriented or a subject-oriented reading. Consider for example the sentence from Keenan & Faltz (1985) with a source PP: *John grabbed the child from the bus*, in which the modifier *from the bus* can be interpreted as either giving the location of the denotation of the subject or of the object.



**4.2.4.4 Ditransitive Verbs (Two animate participants)** The prototypical example under this group of verbs is that of “change of possession” verbs:

- (68)
- a. Mary gave the book to John
  - b. Mary gave John the book
  - c. **es:** María dio el libro a Juan
  - d. **nl:** Marie gaf het boek aan John
  - e. **nl:** Marie gaf John het boek

Other classes of verbs included here are:

- verbs of communication

- (69)
- a. Mary told a story to John
  - b. Mary told John a story

- verbs describing perlocutionary effects

- (70)
- a. Mary warned John of the consequences
  - b. **fr:** Marie a averti Jean des conséquences
  - c. **es:** María advirtió a Juan de las consecuencias
  - d. **de:** Maria hat Johannes vor den Folgen gewarnt

- (71)
- a. John reminded her of the consequences
  - b. **fr:** Jean lui a rappelé les conséquences

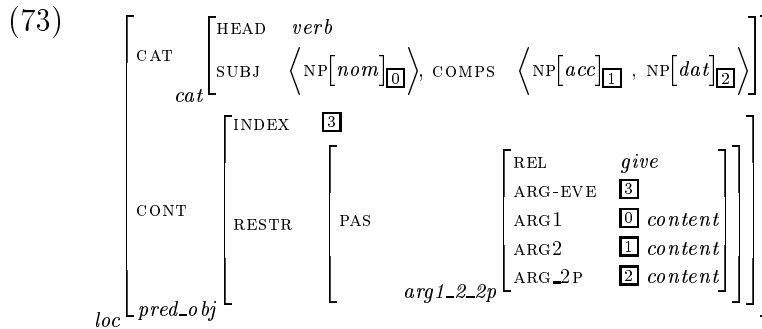
- other verbs with two animate participants

- (72)
- a. Mary blamed the failure on John
  - b. Mary blamed John for the failure
  - c. John deprived Mary of access to the SUN
  - d. John congratulated Mary on her success



- e. John charged Mary with negligence
- f. Mary forgave John (for) his bad behaviour

All these cases can receive a standard lexical representation which assigns the arguments the argument labels ARG1, ARG2, ARG\_2P, as shown in the following entry for *give*:



Note that these verbs allow for different alternations: double object constructions, passivisation and impersonalisation, second animate participant optionality, and ARG2 optionality. The first two alternations listed are rather well known and it is possible to envisage a general treatment that establishes the lexical links between the two classes of lexical entries. The situation is not the same with respect to the optionality factors with respect to ARG2s and ARG\_2Ps; a careful study of these will probably indicate that the generalisation of ARG\_2P is not a correct one, that is to say, it may show that a more fine-grained set of distinctions is more explicative than the one used in the current framework. This however cannot be predicted at the moment, since the study of the impact of thematic roles on PAS is not an established one and many aspects are still unknown.

**4.2.4.5 Ditransitive Verbs (Two inanimate participants)** A set of verbs is treated under this label that have in common that their PAS involves two essentially inanimate entities. This class is exemplified by the following sentences:

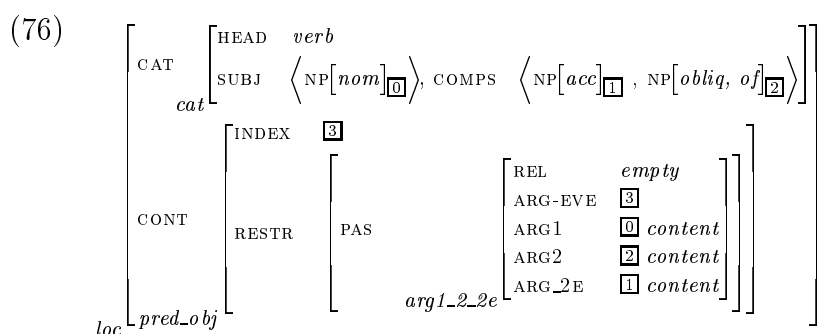
- (74)
- a. John emptied / drained the tank of petrol
  - b. John stripped the tree of bark
  - c. John sprinkled the table with cocaine
  - d. John loaded the cart with grass

Many of these verbs allow (in some languages, but not in others —this is highly language specific) for an alternation with an argument structure which can be easily coded with the ARG\_ADJ label:

- (75)
- a. John emptied / drained the petrol from the tank
  - b. John stripped the bark from the tree

- c. John sprinkled the cocaine on the table
- d. John loaded the grass into the cart

However in the first set of sentences there is no way of using the ARG\_ADJ label to type the prepositional argument. The indirect object must be labelled ARG2 since it is the theme, but the direct object cannot obtain an ARG\_ADJ label. This is the argument typed as ARG\_2E. Of course, the lexical representation of such verbs is quite straightforward, as exemplified in the following for the verb *empty* in this syntactic use (compare it with the one included in section 4.2.4.1):



Two sorts of alterations have to be considered here: passivisation and impersonalisation, and the alternation between the ARG1\_2\_ADJ and ARG1\_2\_2E frames. The first one follows the general lines common to many classes of verbs. The second one, however, is specific to this type of verbs. As noted they seem to be highly language specific. In addition, the two PAS usually imply differences in *Aktionsart*, which can be easily seen if one considers the behaviour of the following alternate sentences:

- (77) a. John half loaded the cart with the grass  
b. John half loaded the grass into the cart
- (78) a. John half loaded the cart with grass  
b. ?? John half loaded grass into the cart

**4.2.4.6 Raising Verbs** Raising verbs allow for a dative perceiver complement, which has been labelled ARG\_PE. Although it must be noted that there is a certain uneasiness in the restricted use of this label (given that dative perceivers occur with verbs like *it: piacere*, which in the current proposal are treated as ARG2), there is no major problem in representing them as we suggest.

A further point to note is that the ARG2 of raising predicates is a *psoa*, and that it can be syntactically realised either as a complete (saturated) element (79) or as an incomplete (unsaturated) element which is only saturated by the raised complement (80):

- (79) a. It seems to Peter that the effort is worthwhile  
b. It seems that the effort is worthwhile

- (80) a. John seems ill (to me)  
 b. **de:** Johannes scheint mir krank zu sein

The lexical representation of these complements is exemplified in section 4.3.1.4 on raising predicates.

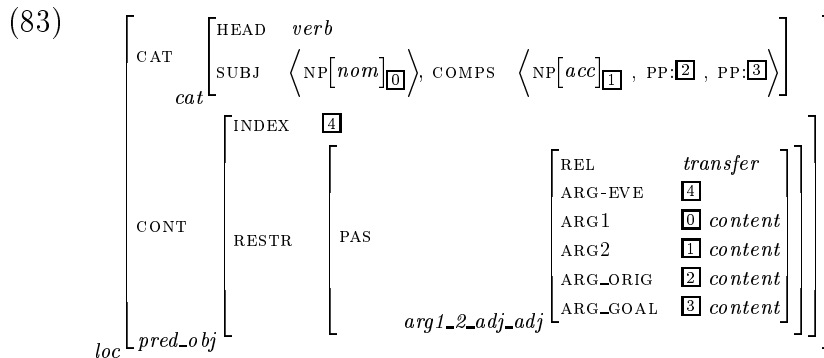
### 4.2.5 Tetravalent Verbs

Two prototypical cases of tetravalent verbs generally considered are “transfer” verbs and “commercial transaction” verbs:

- (81) John transferred his money from one bank to another  
 (82) a. Mary paid John 25 pounds for the book  
 b. John sold Mary the book for 25 pounds  
 c. Mary bought the book from John for 25 pounds

The first class of verbs involves only one animate participant and two locationals (labelled with the corresponding ARG\_ADJs, namely ARG\_ORIG and ARG\_GOAL). The second one however involves two animate participants and a measure argument (which is labelled ARG\_MEASURE).

See a lexical representation for *transfer* exemplifying this kind of verbs:



## 4.3 Clausal Arguments to Verbs

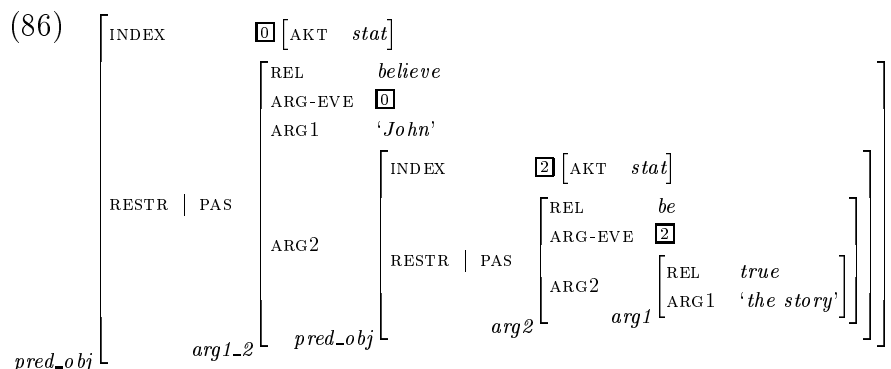
### 4.3.1 Divalent Verbs

**4.3.1.1 Finite Complements** Finite complements to verbs are assigned the argument labels in much the same way as the phrasal ones are. In many contexts the sentential complement alternates with NPs:

- (84) a. John believes the story  
 b. John believes that the story is true

- (85) a. **fr:** Jean croit l’histoire  
 b. **fr:** Jean croit que l’histoire est vraie

Obviously complementisers are not represented at the level of PAS. Thus the partial PAS representation for the sentence (84b) above would be:



Finite complements must be marked with a complementiser, which can cooccur (in some languages has to cooccur) with the strongly bound preposition that marks the complement:

- (87) a. **fr:** La commission consent à ce que les délégués viennent  
 (The commission agrees that the delegates should come)  
 b. **es:** Juan insistió en que María se fuera  
 (Juan insisted that María leave)

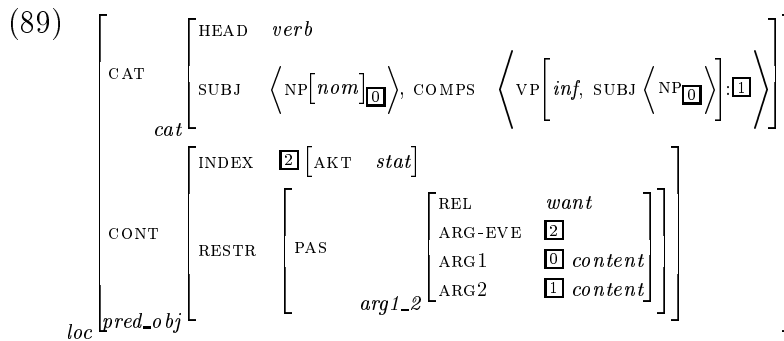
This clearly means that in the cases of languages allowing for such “complex” marking of finite sentences, a distinction is to be drawn between the complementiser and the strongly bound preposition. We suggest that the former be treated as in HPSG and that the latter assign an oblique case with a further specification of PFORM.<sup>16</sup>

**4.3.1.2 Non-finite Complements: Subject Control** This kind of complement is shown in the following example:

- (88) Mary wanted to leave

The following lexical representation for *want* would account for the control of the ARG1 of the non-finite clause:

<sup>16</sup>For a more complete proposal, see Badia (to appear).

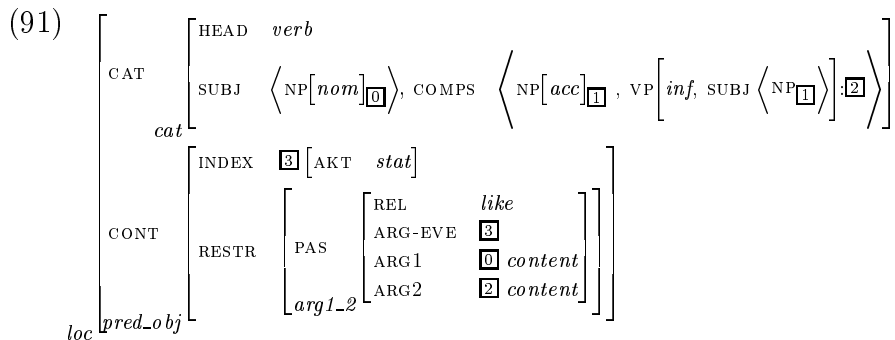


Note that the control of the external argument is realised in a standard way via structure sharing. Note also that some modal verbs can be treated as control verbs and consequently obtain a representation which is identical to this one (see subsection 4.4).

**4.3.1.3 Non-finite Complements: Object Raising** In some cases, an athematic object appears in the main clause and has to be “lowered” into the external argument position of the non-finite clause:

(90) Mary likes his men to cook

The corresponding lexical representation of *like* would be the following:

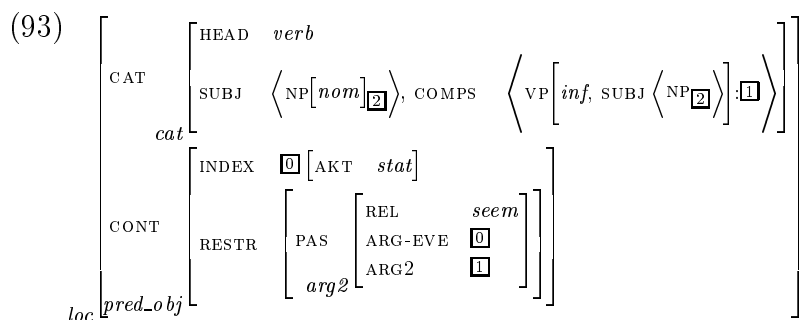


Note that here as well the standard way is used of dealing with control in FS systems.

**4.3.1.4 Non-finite Complements: Raising to Subject** As already commented, athematic subjects are treated as arguments of the embedded clause. Thus the lexical representation for *seem* in the sentence

(92) Mary seemed to be happy (to John)

would be like the following:



Note that a third element in the subcat list should be admitted if the lexical entry should apply to sentences with a dative perceiver. Note also that some modal verbs can be treated as raising verbs and consequently obtain a representation which is identical to this one (see subsection 4.4).

**4.3.1.5 Non-finite Complements: Small Clause** Small clauses admit two possible analyses. A clausal analysis would treat the postnominal elements as some sort of reduced clause. Consider for example the following two sentences:

- (94) a. Mary considered John very mean  
 b. Mary considered John to be a fool

The intuition is that the first construction appears semantically similar to the second one. To take the similarity to the last consequences, such an analysis would require the insertion of some sort of dummy copula. This option however is not open to us given the formalism we are using.

A second possibility is to treat the secondary predicate as an ARG\_AO, so that the verb *consider* in this usage is regarded as trivalent. Under this point of view the adjective governs the noun of which it is the semantic predicator. As already mentioned in the corresponding sections we propose to treat these arguments as attributes to the object and according to what we said in section 4.2.4.3 we adopt the view that constituents appearing in this context are marked as PRD + and are given an external argument the content of which is structure shared with that of the ARG2.

### 4.3.2 Trivalent Verbs

**4.3.2.1 Finite Complements** Here again the finite complement is treated as a phrasal one. The following examples show that there is no difference in this sort of complements be they attached to divalent or to trivalent verbs.

- (95) a. Mary promised John that she would leave  
 b. Mary conceded (to John) that she was wrong  
 c. Mary told John that she would leave

d. Mary concluded from John's behaviour that he was drunk

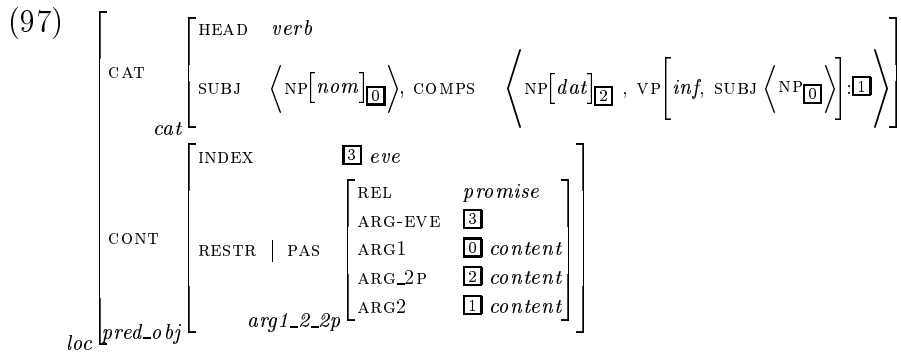
See the note on the marking of such complements in section 4.3.1.1 on finite complements to divalent verbs.

**4.3.2.2 Non Finite Complements: Subject Control** Of course control of non-finite sentential complements is not restricted to divalent verbs. In the case of trivalent verbs, control of that argument can be performed either by the subject of the main clause or by the indirect object.

Here we exemplify the first of these cases:

(96) Mary promised John to leave

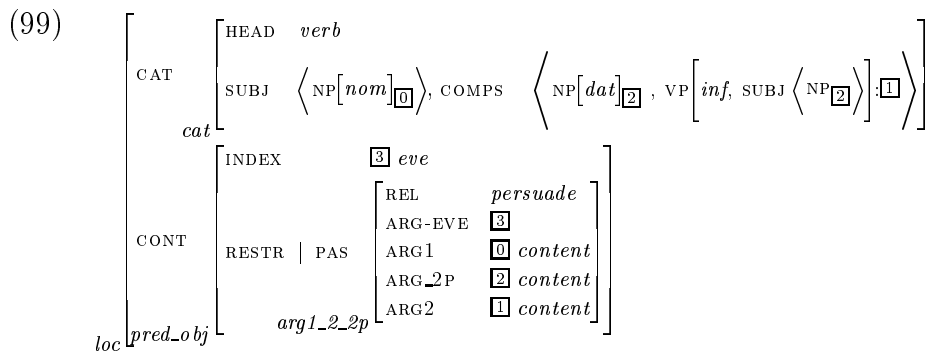
And this is the simplified lexical representation of *promise*:



**4.3.2.3 Non-finite Complements: Indirect Object Control** As an example, consider:

(98) Mary persuaded John to leave

And this is the simplified lexical representation of *persuade*:



## 4.4 A Note on Modal Verbs

Among the verbs with clausal arguments there are the so called modals. Although we are not going to propose a semantic treatment for them (this would bring us quite apart from the main topic of this paper), we certainly should make sure that our PAS theory covers the basic, structural facts about modals. Even if they behave in particular ways, modals clearly belong to the class of verbs having a clausal, non-finite argument. Therefore they are expected to have one of the PAS verbs with clausal arguments actually have.

The two basic possibilities for treating them are as raising or as control verbs. Sometimes it is claimed (e.g., in some older versions of the Eurotra Reference Manual –version 6.1) that epistemic modals are best treated as raising whereas deontic modals should be represented as control verbs. However the borderline between the two classes of verbs is very often not clear and there certainly are mismatches among relatively close languages (such as the Western European ones, which are under the focus of our work). On the other hand we may want to adopt a logical view of modals, just treating them as operators; in this case the most appropriate PAS should be that of a raising predicate with a unique argument labelled ARG2.

Thus we adopt the view (which was also taken in the last version of the Eurotra Reference Manual, 1992) that modals are generally treated as subject raising verbs. Nonetheless a few of them are probably best treated as verbs of control (such as volitional *will* and the ability reading of *can*). We do not provide a specific exemplification for them, but simply point out to the sections where the two classes of verbs were discussed: subject raising verbs have been exemplified in subsection 4.3.1.4, and the control verbs have been shown in subsection 4.3.1.2.

## 5 Exemplification: The Adjectival Signs

### 5.1 Introduction

Adjectives in predicative positions will be treated as having all of them an external argument (labelled ARG1). This implies obviously that this approach is not intended to apply to some classes of adjectives, among which the following have to be included:

- adjectives that belong to the specifier system of the NP: like *last*, *numerous*, and *so on*.

- (100) a. the last book in the row  
b. those numerous friends of yours

- relational adjectives: like *nuclear*, **fr**: *énergétique*, **es**: *mundial*, ...

- (101) a. nuclear physicist



- b. **fr:** consommation énergétique  
(energy consumption)
  - c. **es:** comercio mundial  
(world trade)
- modal adjectives, like *alleged*, *former*, *fake*, ...
- (102)
- a. an alleged novel writer
  - b. the former president of the club
  - c. a fake Bizantine manuscript world trade

In all of these cases the adjective cannot appear in predicative position, i.e., as attribute to the subject or to the object, or as the secondary predicate in “complex predication” structures (small clauses, depictive...):

- (103)
- a. \*the book is last
  - b. \*those friends of yours are numerous
- (104)
- a. \*this physicist is nuclear
  - b. **fr:** \*la consommation est énergétique
- (105)
- a. \*this novel writer is alleged
  - b. \*the president of the club is former
  - c. \*this Bizantine manuscript is fake

Thus the approach to adjectives adopted in this section applies only to the adjectives that can be involved in the following constructions: attribute of subject, attribute of object, resultatives, depictives, free adjuncts and absolutives (which amount to the set of constructions in which adjectives —and other constituents— can be said to be predicative). Among these a determining group is that of the adjectives that allow for raising and extraposition constructions.

## 5.2 Standard, Non-valent Adjectives

In this section we only deal with adjectives in predicative positions; we discuss the representations of adjectives in attributive positions, and the relation between the two classes of uses (and their lexical entries), in section 8 on adjuncts. Thus we restrict the present discussion to the central predicative uses of adjectives.

All adjectives in predicative positions are marked as PRD + and have at least a member in their subcategorisation lists (actually in the SUBJ list), which is the external argument. Thus the simplified lexical entry of *ill* would be the following:

$$(106) \quad \left[ \begin{array}{l} \text{CAT} \\ \text{cat} \\ \text{CONT} \\ \text{arg1} \\ \text{loc} \end{array} \left[ \begin{array}{l} \text{HEAD } \textit{adj}^L \text{ [PRD +]} \\ \text{SUBJ } \langle \text{NP}_{\boxed{0}} \rangle, \text{ COMPS } \langle \rangle \\ \text{REL } \textit{ill} \\ \text{ARG1 } \boxed{0} \textit{ content} \end{array} \right] \right]$$

When the information corresponding to this entry is combined with the one appearing in the COMPS list of a raising verb like *be* (see section 4.2.3.3 for a simplified lexical representation), the corresponding VP (i.e., *be ill*) has the information that results from taking the content of its subject as the content of the argument to *ill*, as shown in section 4.3.1.4 on raising to subject structures.

A very similar move occurs when the adjective marked PRD + appears in other subcategorised for positions: basically, attribute of the object and resultatives. In this case however the main predicate is considered a control predicate (and not a raising one), since the object of the main clause controls the “external argument” of the adjective. For example, when *ill* appears as an attribute of the object of a verb like *believe*, as in (107), the lexical entry for *believe* structure shares the content of its ARG2 with the content of the ARG1 of *ill*.

(107) John believed Mary ill

### 5.3 Valent Adjectives

There are adjectives that have arguments other than this external one:

- adjectives with a prepositional argument

- (108) a. He is fond of children  
 b. He is interested in Medieval music  
 c. **es:** El profesor está contento de sus alumnos  
 (The teacher is satisfied with his pupils)

- adjectives with finite clausal complements

- (109) a. John is happy that Mary has completed the course  
 b. John is fearful that the baby should wake up

- object raising adjectives

- (110) a. The violin is difficult for Mary to play  
 b. For John, this is easy to do

- subject raising adjectives

- (111) a. It is crazy for John to do this  
 b. John is crazy to do this  
 c. Europe is likely to fall

- control adjectives

- (112) a. Bill is eager to win  
 b. Bill is eager for Mary to win

In all of these cases the treatment of the adjectival complement is parallel to that of the verbal ones. Note that in all of these cases, when the adjective appears in a predicative position (and thus it is marked as PRD +) its SUBJ list contains the external argument and the proper argument of the adjective is contained in the COMPS list.

It is worth noting as well that in the current approach we treat *to be* as a raising verb in all occasions.

To conclude, in the following we propose a few lexical representations for these types of adjectives:

- for *fond (of)*:

$$(113) \left[ \begin{array}{l} \text{CAT} \left[ \begin{array}{l} \text{HEAD } \textit{adj} [\text{PRD } +] \\ \text{SUBJ } \langle \text{NP}_{[0]} \rangle, \text{COMPS } \langle \text{NP}[\textit{obliq}; \textit{of}]_{[1]} \rangle \end{array} \right] \\ \textit{cat} \left[ \begin{array}{l} \text{REL } \textit{fond} \\ \text{ARG1 } [0] \textit{content} \\ \text{ARG2 } [1] \textit{content} \end{array} \right] \\ \text{CONT} \\ \textit{arg1-2} \end{array} \right]_{\textit{loc}}$$

- for *happy* with finite sentential complement:

$$(114) \left[ \begin{array}{l} \text{CAT} \left[ \begin{array}{l} \text{HEAD } \textit{adj} [\text{PRD } +] \\ \text{SUBJ } \langle \text{NP}_{[0]} \rangle, \text{COMPS } \langle \text{s}[\textit{fin}; \textit{that}]_{[1]} \rangle \end{array} \right] \\ \textit{cat} \left[ \begin{array}{l} \text{REL } \textit{happy} \\ \text{ARG1 } [0] \textit{content} \\ \text{ARG2 } [1] \textit{content} \end{array} \right] \\ \text{CONT} \\ \textit{arg1-2} \end{array} \right]_{\textit{loc}}$$

- for *likely* in a subject raising structure:

$$(115) \left[ \begin{array}{l} \text{CAT} \left[ \begin{array}{l} \text{HEAD } \textit{adj} [\text{PRD } +] \\ \text{SUBJ } \langle [2] \text{NP}_{[0]} \rangle, \text{COMPS } \langle \text{VP}[\textit{infin} \langle \text{SUBJ } [2] \rangle]_{[1]} \rangle \end{array} \right] \\ \textit{cat} \left[ \begin{array}{l} \text{REL } \textit{likely} \\ \text{ARG2 } [1] \textit{content} \end{array} \right] \\ \text{CONT} \\ \textit{arg2} \end{array} \right]_{\textit{loc}}$$

- for *eager*:

$$(116) \left[ \begin{array}{l} \text{CAT} \\ \text{CONT} \\ \text{loc} \end{array} \left[ \begin{array}{l} \text{HEAD } \textit{adj} \left[ \begin{array}{l} \text{PRD} \\ + \end{array} \right] \\ \text{SUBJ } \langle \text{NP} \left[ \begin{array}{l} \text{2} \\ \text{0} \end{array} \right] \rangle, \text{COMPS } \langle \text{VP} \left[ \begin{array}{l} \textit{inf}, \text{SUBJ} \\ \langle \text{2} \rangle \left[ \begin{array}{l} \text{1} \end{array} \right] \end{array} \right] \rangle \rangle \\ \text{REL } \textit{eager} \\ \text{ARG1 } \left[ \begin{array}{l} \text{0} \\ \textit{content} \end{array} \right] \\ \text{ARG2 } \left[ \begin{array}{l} \text{1} \\ \textit{content} \end{array} \right] \end{array} \right] \right]$$

## 6 Exemplification: The Nominal Signs

### 6.1 Introduction

Nominal signs can be classified according to whether they may have arguments or not; if they can we call them predicative, and if they cannot, non-predicative. This property of nouns does not relate in an absolute way with their morphologic properties; that is to say, nouns derived from predicates can be predicative and non-predicative, and similarly for non-derived nouns. So these two subdivisions of nouns really crossclassify them into four groups: derived and predicative, derived and non-predicative, simple and predicative, and simple and non-predicative.

Among derived nouns two general classes are usually distinguished: predicate and function nominalisations, which differ in the way in which the denotation of the noun is related to that of the source predicate: in the former case the denotation of the derived noun is directly related to that of the predicate, whereas in the latter case it is directly related to that of one of the arguments of the source predicate. These two classes of derived nouns are exemplified by the following sentences (where (117) contains a predicate nominalisation denoting the process of administration and (118) contains a functional nominalisation denoting those who administrate):

- (117) a. **de:** Die Verwaltung der Behörde ist schwierig  
 b. The administration of the office is difficult
- (118) a. **de:** Die Verwaltung hat diese Vorgehensweise beschlossen  
 b. The administration has decided on this procedure

Note that there are some functional nominalisations (such as *analysis*) the denotation of which cannot be said to be related to the one of its arguments but only to one of the participants in the eventuality related to this kind of predication (as shown in the second of the following pair of sentences, where the denotation of the object NP is something like “the report in which Mary’s analysis of the data resulted”):

- (119) a. John’s analysis of the data took an hour

- b. John lost Mary's analysis of the data

In the following sections we are going to deal in turn with predicative and non predicative nouns, and try to provide for them a characterisation of their argument structure.

## 6.2 Basic Representation for Nominal PAS

As mentioned above (section 2.2), the structure of the content of signs that we have adopted facilitates the representation of nominal PAS, even in the cases in which a unique derived noun has two different interpretations (predicative or funtional). Thus for nouns the basic structure of the sign is as follows:

$$(120) \left[ \begin{array}{l} \text{INDEX} \quad \boxed{0} \\ \text{RESTR} \quad \left[ \begin{array}{l} \text{PAS} \quad \left[ \begin{array}{l} \text{REL} \quad \textit{rel} \\ \text{ARG\_N} \quad \boxed{0} \\ \dots \end{array} \right] \\ \text{SEM\_ADJ} \quad \{ \textit{set of contents} \} \end{array} \right] \end{array} \right]$$

Of course the different classes of nouns have different argument labels and index types, and also their indexes may structure share differently with some argument in the *psoa* type. We propose here a representation for the different classes of nominal signs, so that in the discussion below we do not need to build feature structures.

- simple, argumentless, non-predicative nouns (e.g. *girl*)

$$(121) \left[ \begin{array}{l} \text{INDEX} \quad \boxed{0} \quad \left[ \begin{array}{l} \text{PERS} \quad \textit{3rd} \\ \text{NUM} \quad \textit{sing} \end{array} \right] \\ \text{RESTR} \quad \left[ \begin{array}{l} \text{PAS} \quad \left[ \begin{array}{l} \text{REL} \quad \textit{girl} \\ \text{ARG1} \quad \boxed{0} \end{array} \right] \\ \text{SEM\_ADJ} \quad \{ \quad \quad \} \end{array} \right] \end{array} \right]$$

- simple nouns with an argument (e.g., *size*)

$$(122) \left[ \begin{array}{l} \text{INDEX} \quad \boxed{0} \quad \left[ \begin{array}{l} \text{PERS} \quad \textit{3rd} \\ \text{NUM} \quad \textit{sing} \end{array} \right] \\ \text{RESTR} \quad \left[ \begin{array}{l} \text{PAS} \quad \left[ \begin{array}{l} \text{REL} \quad \textit{size} \\ \text{ARG1} \quad \boxed{0} \\ \text{ARG2} \quad \boxed{1} \end{array} \right] \\ \text{SEM\_ADJ} \quad \{ \quad \quad \} \end{array} \right] \end{array} \right]$$

- predicative nouns with a set of arguments (dynamic reading) (e.g. *administration*)

$$(123) \left[ \begin{array}{l} \text{INDEX } \boxed{0} \left[ \begin{array}{l} \text{PERS } 3rd \\ \text{NUM } sing \end{array} \right] \\ \text{ev\_ind\_ind} \\ \text{RESTR } \left[ \begin{array}{l} \text{PAS} \left[ \begin{array}{l} \text{REL } administration \\ \text{ARG-EVE } \boxed{0} \text{ eve} \\ \text{ARG1 } \boxed{1} \\ \text{ARG2 } \boxed{2} \end{array} \right] \\ \text{SEM\_ADJ } \{ \quad \} \end{array} \right] \end{array} \right]$$

- predicative nouns with a set of arguments (functional reading) (e.g. *management*)

$$(124) \left[ \begin{array}{l} \text{INDEX } \boxed{1} \left[ \begin{array}{l} \text{PERS } 3rd \\ \text{NUM } sing \end{array} \right] \\ \text{ind\_index} \\ \text{RESTR } \left[ \begin{array}{l} \text{PAS} \left[ \begin{array}{l} \text{REL } management \\ \text{ARG-EVE } \boxed{0} \text{ eve} \\ \text{ARG1 } \boxed{1} \text{ ind\_index} \\ \text{ARG2 } \boxed{2} \end{array} \right] \\ \text{SEM\_ADJ } \{ \quad \} \end{array} \right] \end{array} \right]$$

Note that in all cases noun arguments have to be considered optional, so that the syntactic part of the sign has to take this fact into account in order to provide for a realistic approach to them.

## 6.3 Predicative Nouns

### 6.3.1 Deadjectival Nouns

As discussed in section 5 adjectives in predicative position have at least an external argument. This argument is usually inherited by the derived noun as the following example shows:

- (125) a. Triangle A is congruent with triangle B  
 b. the congruence of triangle A with triangle B

Naturally there are lexical differences among languages; and particularly what in one language is expressed by a derived nominal in another may be by a simple noun:

- (126) a. **de:** die Grösse von Giraffen  
 b. the size of giraffes

The German noun *Grösse* is deadjectival and therefore inherits the ARG1 from its source adjective. But the English noun *size* is not derived; however it is going to obtain an argument as well according to the principles to be determined for simple nouns (see section 6.4 below).

These nouns could be represented either as simple nouns with a unique argument (as *size* above) or as predicative nouns in their dynamic reading (as *administration* above).

Note that deadjectival nouns may have complex argument structures (just as stative verbs have). For example, the noun *congruence* alternates a two argument PAS with a one argument PAS (provided that it is in plural). Thus in the typing of these arguments the full type hierarchy of the *psoa* type is available to us.

### 6.3.2 Deverbal Nouns

As mentioned above, deverbal nouns can have two different interpretations: functional or predicative. On the other hand, there are many nouns that are ambiguous with respect to these interpretations; for example in English many nominalisations in *-ion* are ambiguous, as shown by the following example:

- (127) a. The construction of the wall (by the brick-layers) took one day  
b. The architect admired the construction (\*by the brick-layers)

Of course, in every language there are some affixes which yield non-ambiguous nominalisations and there are others that produce ambiguous ones.

**6.3.2.1 Dynamic Deverbal Nominalisations** The argument structure of the source predicate is usually inherited by the dynamic nominalisation (although the arguments become optional). As studied in the literature there is a change in case (and/or preposition) assignment, particularly for the semantically unrestricted arguments (our ARG1 and ARG2 arguments). Nonetheless the source predicate is the appropriate place to look for the PAS possibilities of dynamic nominalisations.

All of these nominalisations obtain a representation in which there is an eventive argument, the value of which is structure shared with the value of the index, thus indicating that the nominal is an eventuality denoting one. For an example, see the representation of the dynamic reading of *administration* above.

**6.3.2.2 Functional Deverbal Nominalisations** Functional nominalisations are peculiar in that they usually have arguments and yet they denote individuals (i.e., some participant in the eventuality denoted by the source predicate). Indeed one of the motivations for structuring the content of the sign in the way we do is the treatment of this kind of nominal. Note however that the so-called “relational nouns” have a very similar structure (see subsection 6.4.2.2 below).

In general there are two classes of functional nominalisations: those that denote the individual related to the external argument of the source predicate and those that denote the individual related to its internal argument, to which we refer as ARG1 and ARG2 functional nominalisations respectively.

The most usual class of ARG1 nominalisations are the agentive ones, i.e., those in which the nominal denotes the agent of an eventuality of the kind usually denoted by the source predicate (*manegement* above is an example). There are however nominalisations, such as *opener*, *detector*..., which do not denote the agent but the argument by means of which the action is performed. This kind of nominals usually allow (provided, of course, that they are nominalisations of transitive verbs) for an argument (which corresponds to the second argument of the source predicate). Consequently in their PAS there are going to be the external argument (ARG1) which binds the value of the index and the internal argument (which may be realised by a complement of the noun). The example of functional deverbal noun above (*manegement*) actually is an ARG1 nominalisation.

The other type of functional deverbal nouns (i.e. ARG2 nominalisations) usually cannot have arguments. In these cases the PAS of the nominal is quite simple, since it only has an ARG2, which binds the index.

There is a small class of derived nouns which can still have an internal argument in spite of being result nominalisations. They derive from the source predicates that are sometimes called “redescription” predicates (Rigler, 1988), which are those verbs that denote a process or an event the result of which is not expressed by any of its arguments. This is so because the internal argument (i.e., the *theme*) undergoes the process or event but still remains intact and the activity ends when a different entity has been created. Examples of such predicates are *translate*, *copy*, *analyse*... Consider the following sentence in which the whole of the frame of *translate* is realised:

(128) John translated the last novel by J. Fowels into German

The argument structure of this verb, as used in the example, is *arg1\_2\_goal*. Now what is interesting is that even if all these complements are present in a nominalisation it can still be a result nominalisation:

(129) I haven't read John's translation of the last novel by J. Fowels into German

Here all arguments remain and the nominal phrase clearly does not denote the process of translating the novel, but the resulting product. In these cases, the PAS inherited from the verb is not adequate to represent the one of the nominalisation, because it lacks the argument that should bind the index. Thus for these cases we propose a PAS in which an extra argument is present (which we call ARG-RES), as shown in the following *content* type for *translation*:

(130) 
$$\left[ \begin{array}{l} \text{INDEX} \quad \boxed{4} \textit{ ind} \\ \\ \text{RESTR} \quad \left[ \begin{array}{l} \text{PAS} \quad \left[ \begin{array}{l} \text{REL} \quad \textit{ translation} \\ \text{ARG-EVE} \quad \boxed{0} \textit{ eve} \\ \text{ARG1} \quad \boxed{1} \\ \text{ARG2} \quad \boxed{2} \\ \text{ARG\_GOAL} \quad \boxed{3} \\ \text{ARG-RES} \quad \boxed{4} \end{array} \right] \\ \text{SEM\_ADJ} \quad \{ \quad \} \end{array} \right] \end{array} \right]$$



Note that with this *content* characterisation in the lexical entry we can introduce all arguments of the noun and still clearly represent the fact that the individual denoted by the noun is not any participant included in the original PAS.

## 6.4 Simple Nouns

Simple nouns are non derived nouns. For our purposes there are two main classes of simple nouns: abstract and concrete ones. In the Eurotra Reference Manual (1990), only abstract simple nouns were said to project an argument structure. As we shall see this restriction need not be kept; therefore we are going to deal with the two classes of nouns separately.

### 6.4.1 Abstract Simple Nouns

Frame bearing abstract nouns impose selectional restrictions on their arguments. For example, “duration” is an attribute of events and processes, “width” is an attribute of individuals, and “location” is an attribute of events, processes and individuals.

Following the Eurotra Reference Manual (1990), we classify these nouns in the following groups:

- properties (or attributes)

The NPs headed by these nouns can be described as denoting a property (or attribute: quality, state, internal relation...) of the entity denoted by its argument. These nouns head NPs of one of the following forms:

- (131) a. the  $N_1$  of Det  $N_2$   
 b.  $N_2$ 's  $N_1$   
 c. the  $N_1$  which Det  $N_2$  has
- (132) a. the colour of the house  
 b. the house's colour  
 c. the colour which the house has

- partitions (or partitives)

The NPs headed by these nouns can be described as denoting a non-specific subpart or quantity of the entity denoted by its argument. These nouns head NPs of the following forms:

- (133) a. Det<sub>1</sub>  $N_1$  of Det<sub>2</sub>  $N_2$   
 b. \*  $N_2$ 's  $N_1$
- (134) a. the rest of the people

b. \*the people's rest

- sets (or collectives)

The NPs headed by these nouns can be described as denoting a set of the entities denoted by their argument. These nouns head NPs of the following forms:

- (135) a. Det  $N_1$  of *the/0/quant*  $N_2$ [plu]  
 b. \*  $N_2$ 's  $N_1$   
 c. Det  $N_1$  *which consists of the/0/quant*  $N_2$ [plu]  
 d.  $N_3$ 's  $N_1$  of *the/0/quant*  $N_2$ [plu]  
 e.  $N_3$  *made a*  $N_1$  of *the*  $N_2$
- (136) a. the/a list of the/0/10 students  
 b. \*the students' list  
 c. the/a list which consists of the/0/10 students  
 d. the new professor's list of the/0/10 students  
 e. the new professor made a list of the students

For all these nominals we foresee a PAS type with two arguments: an ARG1 which is coindexed with the INDEX value, and an ARG2 which consists in the complement of the abstract noun. In this way these nominals obtain a similar PAS to the one given to the deadjectival nouns; this of course is fully adequate since there are many pairs of abstract nouns (of different languages) which are deadjectival in one of them and just simple nouns in the other (as observed above with respect to **de**: *Grösse* and *size*).

#### 6.4.2 Concrete Simple Nouns

Concrete simple nouns characteristically denote individuals.<sup>17</sup> Most of them do not have arguments. But there is a small group of them that have arguments. In the following we briefly discuss the two groups.

**6.4.2.1 Non-argument Constructions** There is a set of constructions which have the form indicated in (137) and nonetheless do not establish a head-argument relation.

- (137) Det<sub>1</sub>  $N_1$  of Det<sub>2</sub>  $N_2$

In these cases the noun head of the construction (i.e.,  $N_1$ ) does not express selectional restrictions on  $N_2$ . Possession is the most characteristic of these modifier relations.

This is a tentative list of such constructions:

---

<sup>17</sup>This is an oversimplification: mass nouns do not denote individuals but a mass, or portions of it.

- possession
  - (138) a. John's house
  - b. the house of John's
- constituency
  - (139) a. a house of stone
  - b. a table of wood
- identity
  - (140) a. the problem of funding research
  - b. the field of information technology
- nominal apposition
  - (141) a. the city of Tarragona
  - b. the region of Bretagne
- kind or characteristics
  - (142) a. a man of honour
  - b. a man of courage

**6.4.2.2 Argument Constructions** There are some individual denoting nominals which can have arguments. These are nominals that only obtain their denotation (when it is a concrete one) through the denotation of their arguments; that is to say, they can be seen as expressing a relation between two different individuals (the one denoted by the argument, and the one denoted by the whole construction). These nouns are often called “relational”.

Consider the NP:

(143) a friend of John's

The denotation of the whole NP can only be accessed if we have the denotation of the argument; that is to say, it is only possible to identify the individuals being denoted by the NP if the denotation of *John* is already known.

In addition, in these cases the head noun expresses selectional restrictions on its argument.

These nouns have then an ARG1 which is coindexed with the INDEX and an ARG2 which has the content of their argument as value.

Some picture and semiotic nouns behave in a similar way; they have an argument that denotes the individual of which the denotation of the whole NP is a picture or a description:

- (144) a. a portrait of the president  
 b. a description of the valley

Finally, some simple nouns may have directional complements, which can be easily expressed by the ARG\_ADJs of our PAS theory:

- (145) a. the road to Manresa  
 b. the train to Munic  
 c. the letter to the Comission  
 d. the memorandum from the Chairman

## 7 Exemplification: The Prepositions

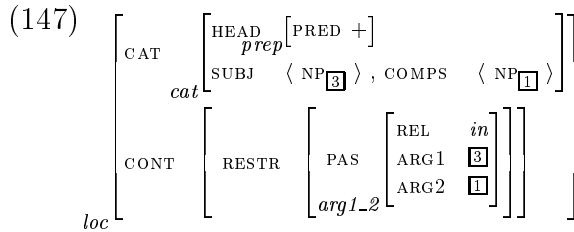
As mentioned in section 2.4 we distinguish between two classes of prepositions (strongly bound and weakly ones) and PPs. As we saw there, the first ones are given a representation which is basically that of a nominal.

Weakly bound prepositions however are considered the head of the construction and therefore are given a representation in which the semantic contribution of the preposition is apparent. A PP with a weakly bound preposition subcategorises at least for a NP complement. The content of this subcategorised complement is the value of the ARG2 feature in the PAS. Note however that, as we said in section 2.4, all PPs headed by semantically full prepositions have an ARG1 as well. There are nevertheless for weakly bound prepositions three possible ways to receive an ARG1 depending on the position they appear in. Let us briefly consider the three possibilities with the respective lexical entries for the preposition *in*.

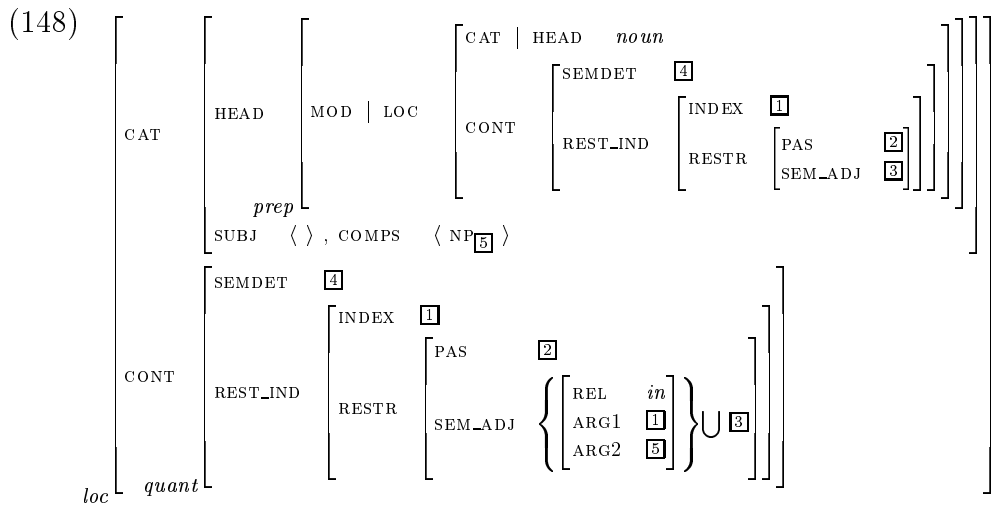
In an ARG\_ADJ position the PP introduces an index of type locative, and their ARG1 is coindexed with that index. Thus the corresponding lexical entry for the preposition is going to be as follows:

$$(146) \left[ \begin{array}{l} \text{CAT} \\ \text{CONT} \end{array} \left[ \begin{array}{l} \text{cat} \\ \text{RESTR} \end{array} \left[ \begin{array}{l} \left[ \begin{array}{l} \text{HEAD } prep \\ \text{SUBJ } \langle \rangle, \text{ COMPS } \langle \text{NP}_{\mathbb{1}} \rangle \end{array} \right] \\ \text{INDEX } \mathbb{3} \text{ loc\_index} \\ \text{PAS } \left[ \begin{array}{l} \text{REL } in \\ \text{ARG1 } \mathbb{3} \\ \text{ARG2 } \mathbb{1} \end{array} \right] \end{array} \right] \right] \right]$$

Secondly, weakly bound PPs may appear in predicative position as attribute to the subject or to the object. In the first case the ARG1 is occupied by the content of the raised NP and in the second one it is structure shared with the content of the controller NP. See the following lexical entry for the preposition *in* as heading weakly bound PPs in predicative positions:



Finally, as a modifier, the ARG1 of the semantically full PPs is structure shared with the index introduced by the head of the construction (the noun or the verb). Therefore, the lexical entry of the preposition in these cases is going to be as follows:



## 8 Exemplification: The Adjuncts

### 8.1 Introduction

One of the big problems in relation to adjuncts is to determine the boundary between them and arguments. Of course there are cases that are absolutely clear, as the direct object of verbs or some prepositional complements headed by a strongly bound preposition. However in many cases the distinction is not clear. Consider for example the dative complements in the following sentences.

- (149) a. I gave him a book  
 b. I found him a book

The relation of the predicate with the dative complement is not the same one in the two cases; we can informally say that *him* in the first one obtains a recipient role, whereas in the second example it obtains the benefactive one. This distinction has a syntactic counterpart in many languages. In English when this element is construed as a PP, the preposition appearing in the complement is usually a different one: recipients take *to* and benefactives take *for* (similar distinctions appear in other languages):

- (150) a. I gave a book to him  
b. I found a book for him

There seems to be a distinction between these two classes of complements, in particular when the fact is considered that most verbs can accept a benefactive complement, which is almost always an optional dependent. The former belongs to the core of the predication and the latter to its periphery (Somers, 1987).

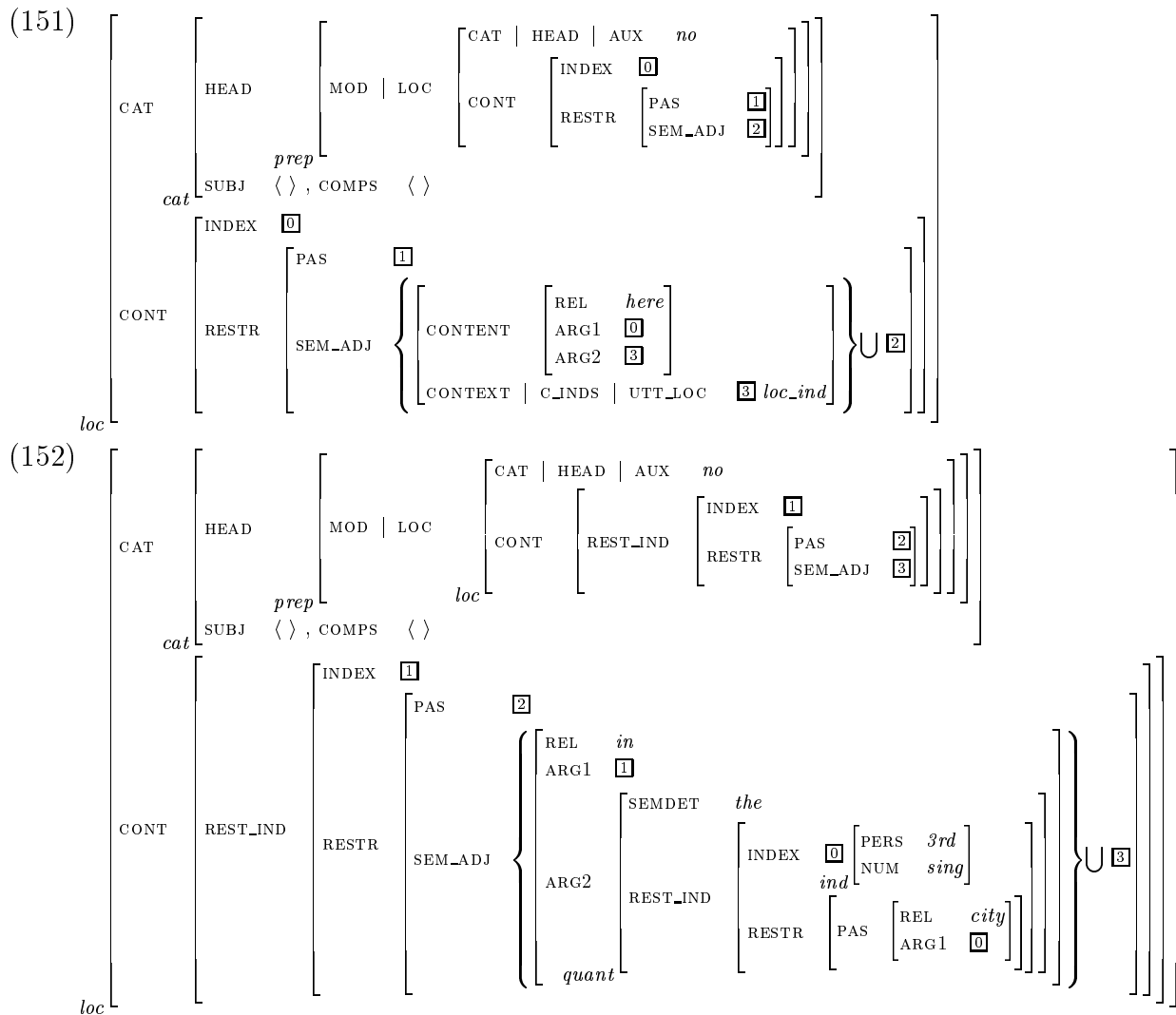
## 8.2 The Representation of Adjuncts

Characteristically adjuncts are not particularly determined by their heads, they are optional, they can be iterated, and their contribution to the content of the phrase is radically different from that of subcategorised for complements. Thus the kind of relation that is established between heads and complements is not adequate to head-adjunct relations: lists are not an adequate way to represent this relation (since they are not flexible enough to deal with optionality and iteration), and the way in which the complements contribute to the content of the phrase cannot be adapted for adjuncts (since their content usually does not fill an argument place of the PAS of the head).

In Pollard & Sag (1987) the idea was presented that heads select their adjuncts, so that there was a HEAD feature in the syntax of the head, called ADJUNCT in which the different classes of adjuncts were contained in a set. This had two different difficulties: firstly, only the syntactic aspects of the adjuncts could be specified, since iteration has to be allowed and therefore both phonological and semantic information could not be unified; and secondly, head selection of adjuncts presented serious difficulties for the calculation of what the content of adjuncts contributes to the content of the whole phrase (especially when there were more than one adjunct in a phrase).

Particularly for the last reason mentioned, Pollard and Sag have reformulated their account of the relation between heads and adjuncts. In Pollard & Sag (1994) they take the opposite view, namely that adjuncts select their heads, on line with what has always been argued in CG. The basic idea is that in a head-adjunct construction the content of the phrase is identical to the content of the adjunct, in which the content of the head has been incorporated via structure sharing. This means that adjuncts can determine the syntactic class of their heads, but also that they can specify the semantic properties they should have. Some adjectives for example can only apply to concrete nouns (e.g., colour adjectives), and some adjuncts can only apply to non-stative predicates (like punctual temporal adverbials). These conditions can easily be stated under this approach.

Let us see, for example, how a PP adjunct and an adverbial would be represented. We assume that no principled distinction is to be done between ADVPS and PPs, and that apart from differences in meaning what distinguishes them is the structure of their content (which is a consequence of their internal structure: for the former there are no complement daughters, whereas for the latter there is one such daughter).

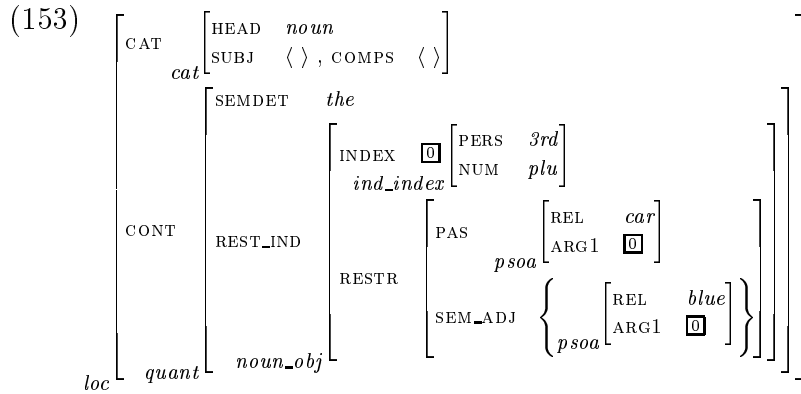


Note that following Pollard & Sag (1994) we have introduced a new feature, MOD, which states the conditions a head has to comply with in order to be modified by that adjunct. Thus, in the representation of the PP *in the city* these conditions are specified as the value of the path SYNSEM | LOC | CAT | HEAD | MOD: thus it demands of its head that it be verbal, and that it be dynamic; and similarly for *here*. As presented in Pollard & Sag (1994), this solution applies equally well to adjuncts to predicative heads and to adjuncts to nominal heads; in fact, the examples presented above are of the former type, whereas those discussed in Pollard and Sag's book are of the latter kind. Thus the two basic classes of adjuncts can be easily implemented in a similar way.

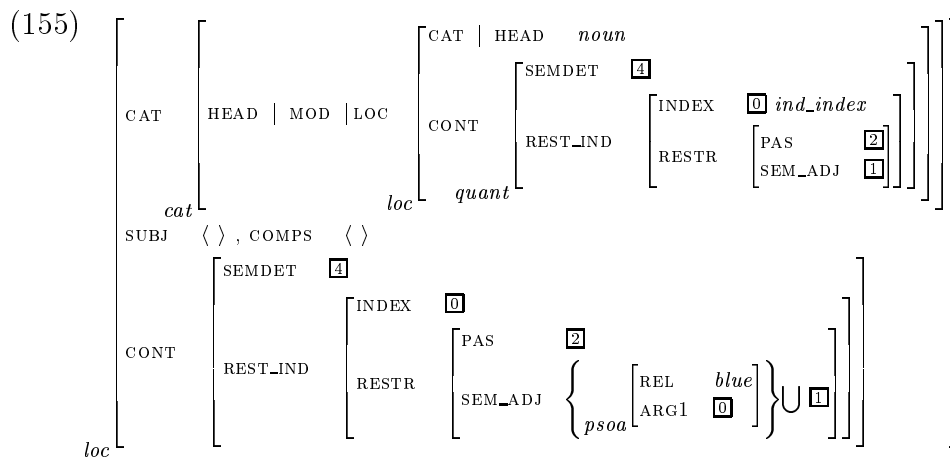
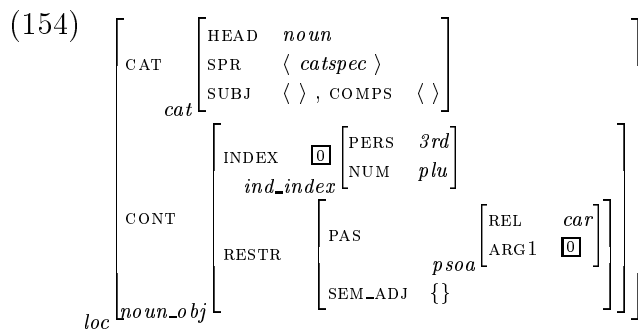
### 8.2.1 The Adjuncts to Nominal Signs

Nominal signs have a set-valued feature in which the restrictions to the parameter are stated. The lexical content of the head noun of the sign is represented as a *psoa* which is the value of the PAS attribute; and the content of all (intersective) adjuncts can be

represented as further restrictions in the restrictions set.<sup>18</sup> In this sense the following is the simplified representation of the nominal sign *the blue cars*:



This representation is achieved by combining the following two lexical representations (for the noun *cars* and the adjective *blue*), together with the one for the determiner (see the chapter on determination in Schmidt, 1996).



Most adjunct adjectives can be treated in this way, or else they can be treated in a way that is a natural extension of this treatment (as proposed by Pollard and Sag for modal

<sup>18</sup>This treatment is obviously not appropriate for non intersective adjuncts, for which the content of the phrase does not result from the intersection between the content of the head noun and the content of the adjunct.



adjectives –see also Arnold & Sadler, 1992). Also different syntactic classes of adjuncts to nominal signs can be treated in a similar way: clear adjunct PPs, relative sentences... can be equally treated as specifying some extra restrictions to the index introduced by the head noun.

Finally it is worth mentioning that most of the constituents that can appear as adjuncts to nominal signs can also appear as predicates in attribute position. In these cases it is worth considering the possibility of establishing links between the two lexical entries, e.g. via lexical rules. How this can be implemented has been shown by Arnold & Sadler, (1992) with respect to the plain intersective adjectives. Their lexical rule, adapted to the current general type system, is the following:

$$(156) \quad \left[ \begin{array}{l} \text{loc} \\ \text{CAT} \\ \text{CONT} \end{array} \left[ \begin{array}{l} \text{HEAD } \textit{adj} [\text{PRD } +] \\ \text{SUBJ } \langle \boxed{0} \rangle, \text{COMPS } \langle \boxed{4} \rangle \\ \text{cat} \\ \boxed{1} \end{array} \right] \right] \Rightarrow \left[ \begin{array}{l} \text{loc} \\ \text{CAT} \\ \text{CONT} \end{array} \left[ \begin{array}{l} \text{HEAD } | \text{MOD } \boxed{0} \\ \text{SUBJ } \langle \rangle, \text{COMPS } \langle \boxed{4} \rangle \\ \text{cat} \\ \text{noun\_obj} \left[ \begin{array}{l} \text{CAT } | \text{HEAD } \textit{noun} \\ \text{CONT} \left[ \begin{array}{l} \text{INDEX } \boxed{5} \textit{index} \\ \text{RESTR } \left[ \begin{array}{l} \text{PAS } \boxed{3} \\ \text{SEM\_ADJ } \boxed{2} \end{array} \right] \end{array} \right] \\ \text{noun\_obj} \left[ \begin{array}{l} \text{INDEX } \boxed{5} \\ \text{RESTR } \left[ \begin{array}{l} \text{PAS } \boxed{3} \\ \text{SEM\_ADJ } \{ \boxed{1} \} \cup \boxed{2} \end{array} \right] \end{array} \right] \end{array} \right] \end{array} \right] \right]$$

### 8.2.2 The Adjuncts to Predicative Signs

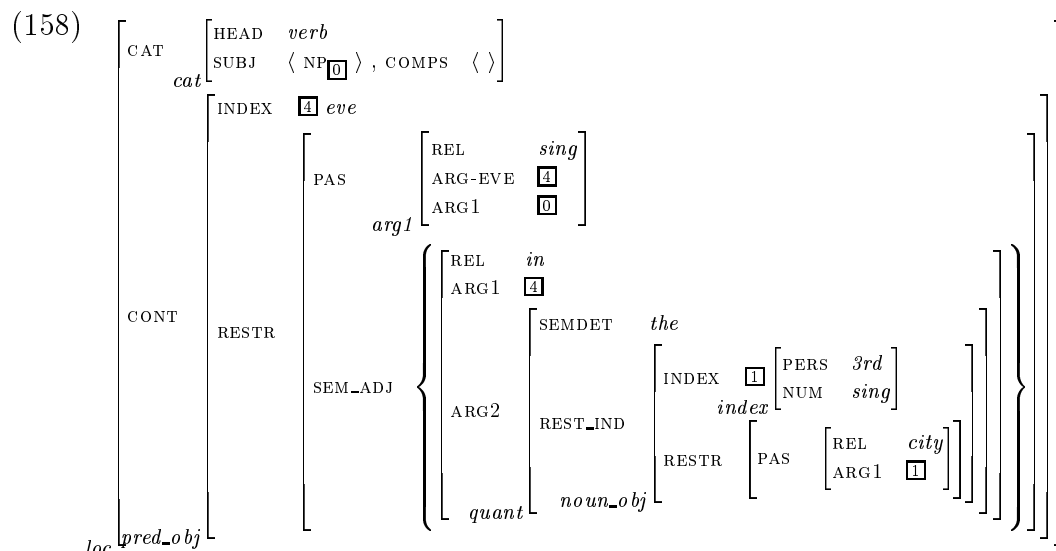
As we have already said, in the present approach adjuncts to predicative signs can be treated in a similar way, because also for predicative signs we have a slot available where to insert the content of adjuncts. As is well known this is one of the main advantages an eventive approach provides. Since we have an index for eventualities in a parallel way as nominal signs introduce an index over individuals, adjuncts can impose restrictions on them in a similar way as they do in nominal signs. This approach takes the view that adjuncts to predications can be viewed as contributing to the phrasal content by specifying additional semantic restrictions to the PAS expressed by the predicative head and its arguments.<sup>19</sup> Recall that the general structure of the content part of the predicative sign is the following:

$$(157) \quad \left[ \begin{array}{l} \text{pred\_obj} \\ \text{INDEX } \boxed{0} \\ \text{RESTR } \left[ \begin{array}{l} \text{PAS } \textit{psoa} \\ \text{SEM\_ADJ } \{ \textit{set\_of\_cont} \} \end{array} \right] \end{array} \right]$$

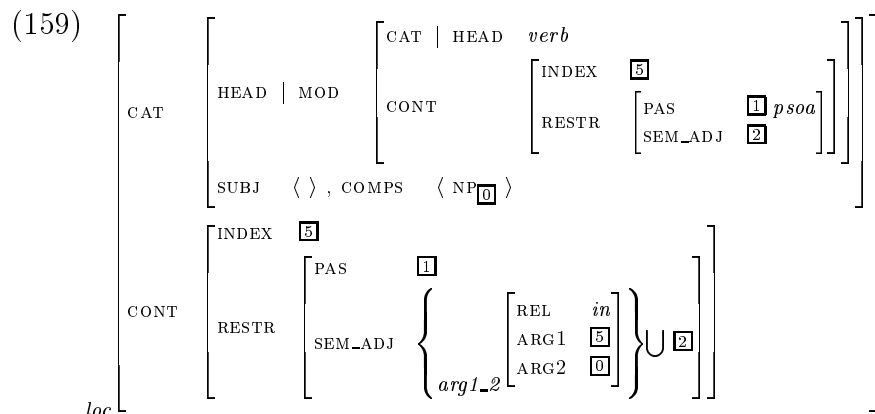
By constructing these structures for the content of phrasal, predicative signs we are adopting an approach to adjuncts to predicates which is very similar to the one that is standard for adjuncts to nominal signs. They can be treated as the semantic heads

<sup>19</sup>A relatively related treatment is to be found in Kasper's paper on the adjuncts in the German *Mittelfeld* (Kasper, 1994), in which a proposal is developed that allows the formulation of flat rules that accept arguments and adjuncts in any order (as it is the case with the German *Mittelfeld*).

of the construction and their contribution is to add a further restriction to the set of restrictions which appear in the content of their (phrasal) sister. Thus the lexical entries of this class of adjuncts are very similar to those of the other class. For example, the (simplified) feature structure corresponding to the phrase *sing in the city* would be the following:



The lexical entry for the preposition *in* in a construction like this one would be following:



Note that this lexical entry in essence does the same job that the lexical entries for adjectives did with respect to nominal adjuncts. There is a feature MOD in which the conditions are specified that the predicative phrase has to comply with in order to be modified by a PP headed by this preposition (which in the previous lexical entry are just the syntactic condition that it be a verbal phrase, and the semantic structuring of the sign which has to be reconstructed in the mother feature structure — of course those conditions can be expanded as needed for the particular lexical entries in realistic grammars). At the same time the feature COMPS states the characteristics of the sign the preposition has to combine with in order to form an adjunct PP of the required sort. Finally, the semantics of the mother node is obtained by letting the semantic parts of

the head daughter be structure shared with that of the mother, except for the value of the *RESTR*, which is the result of the union of the restrictions already present in the head daughter content with the one newly introduced by the adjunct headed by *in*.

A similar representation can be obtained for other classes of adjuncts: adverbial phrases, sentences, and some NPs (e.g., temporal NPs in many languages).

### 8.3 Non-intersective Adjuncts

As is well known not all adjuncts are intersective. This means that the semantically oriented treatment that we have proposed for adjuncts need not be adequate for all classes of adjuncts. As a matter of fact this is so: there are different sorts of adjuncts the semantic interpretation of which cannot be performed on the line we have taken.

This is particularly clear if we consider those adjuncts that can be affected by scope; in all cases in which an adjunct may have scope over another or over another kind of constituent, the inclusion of adjuncts within a set is not going to be semantically appropriate. The work on adjuncts in Categorical Grammar has showed that in many cases adjuncts have to be treated as predicates of their heads, that is to say as operators which take as arguments the categories they modify.

Kasper’s proposal for the treatment of adjuncts within HPSG (Kasper, 1994) characterises two classes of adjuncts (restrictive adjuncts and operators), and treats them in a different way. The treatment he proposes for restrictive adjuncts resembles to a certain extent the one we have followed in the sections above; the basic differences lie in what the modifier restricts (in his approach it restricts a location attribute that is present in all process predications; in ours it restricts the eventuality index of the predication), but in both approaches the content of the modifier is included in a restrictions list.

However Kasper’s proposal for operator modifiers is quite different. He treats them as real semantic predicates, which take as argument the content of their heads. Thus he proposes the following lexical entry for the German causal preposition *wegen* when it subcategorises for an NP (slightly adapted to the current type system):<sup>20</sup>

$$(160) \left[ \begin{array}{l} \text{CAT} \\ \text{CONT} \end{array} \left[ \begin{array}{l} \text{HEAD} \\ \text{SUBJ} \\ \text{COMPS} \end{array} \left[ \begin{array}{l} \text{MOD} \\ \langle \text{NP}_{[2]} \rangle \\ \langle \rangle \end{array} \right] \right] \right]$$

$$\left[ \begin{array}{l} \text{MOD} \\ \text{COMPS} \end{array} \left[ \begin{array}{l} \text{CAT} \mid \text{HEAD} \textit{verb} \\ \text{CONT} \quad \boxed{1} \textit{psoa} \end{array} \right] \right]$$

$$\left[ \begin{array}{l} \text{CAUSE} \\ \text{CAUSED-SOA} \end{array} \left[ \begin{array}{l} \boxed{2} \\ \boxed{1} \end{array} \right] \right]$$

*loc*      *cause-soa*

This treatment could be adapted to our framework, whenever it is clear that the adjunct is correctly interpreted as an operator. In fact there are no technical problems to take

<sup>20</sup>The final proposal he adopts differs slightly from the one reproduced here. We take however this one because in it the basic structure of the sign is more similar to ours. Of course there are still differences, but they can certainly be easily overcome.

this approach.

Such a move however takes us far away from the setting that we defined for our PAS representation. As a matter of fact, such a characterisation of some adjuncts determines the abandonment of the dependency representation. This in itself would not be a problem if the alternative was a clearly defined one. But unfortunately this is not so: a full semantic characterisation of adjuncts has not been achieved yet, and there are many cases which are not clearly understood. Moreover, it must be borne in mind here that it is a decision of our work not to represent scope, that is to say that the deepest level of representation we aim at does not contain any reference to the relative scope of the constituents involved.

So for the time being we would suggest to adopt a more simple approach to modifiers: to treat them all alike, i.e., as members of the restriction set, and to make explicit the relation they establish with their head in the way suggested in next section.

## 8.4 The Labelling of Adjuncts

The distinction between modifiers to predicates and to non-predicates becomes a fundamental distinction (with syntactic import) when one tries to develop an implementation of adjuncts which is coherent with the main-stream linguistic description based on constraints. Adjuncts select their heads, and in most cases the selection applies only to a class of heads (predicative or non predicative).

Given that in the current proposal adjuncts are all introduced in a restrictions set without really elaborating on the semantic relation that they have with whatever is the semantic content of the head of the construction, it seems advisable to type adjuncts in some way.<sup>21</sup> This could be easily done by simply allowing a complex value for the attribute REL, so that it contains both the name of the relation and the type of relation (i.e., the MODSR value –the modifying semantic relation–, in the terminology employed in Eurotra; Durand, 1992), as shown in the following partial representation of the content of the preposition *in* when it heads a locative adjunct to a predicative head:

$$(161) \left[ \text{SEM\_ADJ} \left\{ \begin{array}{l} \text{REL} \left[ \begin{array}{l} \text{REL\_TYPE} \quad \textit{place} \\ \text{REL\_NAME} \quad \textit{in} \end{array} \right] \\ \text{ARG1} \quad \boxed{1} \\ \text{ARG2} \quad \boxed{0} \end{array} \right. \right. \right]$$

Of course the introduction of such a typing implies a lexical distinction between the different senses a preposition (or any other head of an adjunct) may have. This means that after compilation there must be as many different lexical entries as different senses there are; however any TFS formalism provides means to avoid the multiplication of

---

<sup>21</sup> An interesting proposal is the one developed within the Eurotra project; see a description in Durand, 1992.

lexical entries in the grammar writer's bench, which can be easily used in this case; basically simple alternation of values would do in this case.

## Conclusion

In this paper we have shown that a consistent and comprehensive approach to PAS is possible, which takes into account the vast majority of PAS structures. In order to obtain it we have devised a level of description (which ultimately comes from the dependency tradition as applied to Machine Translation) which has enabled us to maintain an adequate depth in our representations: not too surfacy (otherwise we would not have been able to represent in a satisfactory way many of the predicative structures (attributes, prepositions...)), but not too deep either (since otherwise we would not have been able to provide a generally valid approach (we do not have at present a general theory of modification)).

In the particular sections above we have shown by way of exemplification that such an approach is indeed possible and that it can be reasonably levelled for most of the predicative structures present in the languages considered.

It is also important to notice that the work reported here has not been done in isolation, but in conjunction with a lot of other perspectives relevant to natural language processing in typed feature-structure formalisms. For a full understanding of this work we refer the reader to the final report of the project which is at the basis of the work described here (Schmidt, 1996).

## Bibliography

- Allegranza, V., P. Bennett, J. Durand, F. van Eynde, L. Humphreys, P. Schmidt & E. Steiner (1991), "Linguistics for MT: the Eurotra linguistic specifications", in C. Copeland, J. Durand, S. Krawer & B. Maegaard (eds.), *Studies in Machine Translation and Natural Language Processing*, vol. 1. Office for Official Publications of the European Communities, Luxembourg, 15-123.
- Alshawi, H.; D.J. Arnold; R. Backofen; D.M. Carter; J. Lindop; K. Netter; S.G. Pulman; J. Tsujii & H. Uszkoreit (1991), *Eurotra ET/1: Rule Formalism and Virtual Machine design Study (Final report)*. CEC. Luxembourg.
- Arnold, D. & L. Sadler (1992), "Noun-modifying adjectives in HPSG". Working Papers in Language Processing, 15. University of Essex.
- Badia, T. (1992), "Dependency and machine translation", in F. van Eynde (ed.), *Linguistic issues in machine translation*. Pinter Publishers.
- Badia, T. (1994), "The Spanish grammar", in Markantonatou & Sadler (1994a).

- Badia, T. (to appear), "Prepositions in Catalan", in Balari, S. & L. Dini (eds.), *Head-Driven Phrase Structure Grammar in Romance*. CSLI. Stanford.
- Borsley, R. (1989), "An HPSG approach to Welsh", in *Journal of Linguistics*, 25.
- Borsley, R. (1990), "Welsh passives", in Ball, Poppe & Rowland (eds.), *Celtic Linguistics*. Philadelphia and Amsterdam: Benjamin.
- Bresnan, J. (ed.) (1982), *The mental representation of grammatical relations*. MIT Press.
- Carpenter, B. (1992), *The Logic of Typed Feature Structures*. Cambridge: Cambridge University Press.
- Chomsky, N. (1981), *Lectures on government and binding*. Foris.
- Davidson, D. (1967), "The logical form of action sentences", in Rescher, N. (ed.), *The Logic of Decision and Action*. Pittsburgh: University of Pittsburgh Press.
- Durand, J. (1992), "On the translation of prepositions in multilingual MT", in F. van Eynde (ed.), *Linguistic issues in machine translation*. Pinter Publishers.
- Durand, J., P. Bennett, V. Allegranza, F. van Eynde, L. Humphreys, P. Schmidt and E. Steiner (1991), "The Eurotra linguistic specifications: An overview", in *Machine Translation*, 6, pp. 103-147.
- ET-RM 7 (1992), *Eurotra Reference Manual*, 7. Commission of the European Communities, Luxembourg.
- Gebruers, R. (1988), "Valency and MT: Recent developments in the Metal system", in *Proceedings of the second conference of applied natural language processing*, Austin, Texas.
- Halliday, M. (1967-8), "Note on transitivity and theme", *Journal of Linguistics*, 3.
- Halliday, M. (1970), "Language function and language structure", in J. Lyons (ed.), *New horizons in linguistics*. Penguin.
- Hudson, R. (1987), "Zwicky on heads", in *Journal of linguistics*, 23, pp. 109-32.
- Jackendoff, R. (1977), *X-bar syntax: A study of phrase structure*. MIT Press.
- Jackendoff, R. (1983), *Semantics and cognition*. MIT Press.
- Kasper, R. (1994), "Adjuncts in the Mittelfeld", in J. Nerbonne, K. Netter, & C. Pollard (eds.), *German Grammar in HPSG*. CSLI. Stanford.
- Keenan, E. & L. Faltz (1985), *Boolean semantics for natural language*. D. Reidel. Dordrecht.
- Kirschner, Z. (1987), "APAC3-2: An English to Czech MT system", in *Explizite Beschreibung der Sprache und automatische Textbearbeitung*, vol. XIII. Charles University. Prague.
- Markantonatou, S. & L. Sadler (1994a), *Grammatical Formalisms: Issues in Migration*. Luxembourg: Commission of the European Communities.
- Markantonatou, S. & L. Sadler (1994b), "Techniques and devices", in Markantonatou &

Sadler (1994a).

Markantonatou, S. & L. Sadler (1996), “Verb Alternations by Underpsecification”. Department of Language and Linguistics Working Paper No 50, University of Essex.

Markantonatou, S. & L. Sadler (to appear), “Linking Indirect Arguments: Verb Alternations in English” in Godard, D. & J. Jayet (eds.) *Proceedings of CSSP95*.

Perlmutter, D.M. (1983) (ed.) *Studies in Relational Grammar*, 1. University of Chicago Press.

Perlmutter, D.M. & C.G. Rosen (1984) (ed.) *Studies in Relational Grammar*, 2. University of Chicago Press.

Pollard, C. & I. Sag (1987), *Information-based Syntax and Semantics*. Stanford, Calif: CSLI.

Pollard, C. & I. Sag (1994), *Head-Driven Phrase Structure Grammar*. Stanford, Calif: CSLI — The University of Chicago Press.

Rigler, E. (1988), “The semantic classification of deverbal nouns”. Working Papers in Language Processing No 2. University of Essex.

Sanfilippo, A. (1990), *Grammatical Relations, Thematic Roles and Verb Semantics*. PhD Dissertation, University of Edinburgh.

Schmidt, P. (1994), “The German grammar”, in Markantonatou & Sadler (1994a).

Schmidt, P. (ed.) (1996), Final Report. Investigation of linguistic specifications for future industrial standards: The EUROTRA Reference Manual. CEC. Luxembourg. Available at the following URL address: <http://www.iai.uni-sb.de/REFMAN/pap-refm.html>

Schubert, K. (1987), *metataxis: Contrastive dependency syntax for machine translation*. Foris.

Somers, H. (1986), “The need for MT-oriented versions of case and valency in MT”, in *Proceedings of the 11th international conference on Computational Linguistics. Coling'86*. University of Bonn.

Somers, H. (1987), *Valency and Case in Computational Linguistics*. Edinburgh: Edinburgh University Press.

Vauquois, B. (1988), *Bernard Vauquois et la TAO* (ed. part Ch. Boitet). Association Champollion. Grenoble.

Wilkins, W. (ed.) (1988), *Thematic relations*. Academic Press.

Zaenen, A. (1988), “Unaccusatives in Dutch and the syntax-semantics interface”. CSLI Report 123. Stanford University.

Zwicky, A. (1985), “Heads”, in *Journal of linguistics*, 21, pp. 1-29