

**Generating Spanish Clitics  
with  
Constrained Discontinuous Grammars**

**by**

**Charles Grant Brown**

MASc The University of British Columbia 1970  
BASc The University of British Columbia 1967  
DipEngPhys Le Collège Militaire Royal de Saint Jean 1965

A THESIS SUBMITTED IN PARTIAL FULFILLMENT OF  
THE REQUIREMENTS FOR THE DEGREE OF  
Doctor of Philosophy  
in the School  
of  
Computing Science

SIMON FRASER UNIVERSITY  
July 1987

All rights reserved This thesis may not be  
reproduced in whole or in part, by photocopy  
or other means, without the permission of the author.

Copyright © 1987 Charles Grant Brown

## Approval

Name: Charles Grant Brown  
Degree: Doctor of Philosophy  
Title of Thesis: Generating Spanish Clitics with Constrained Discontinuous Grammars

\_\_\_\_\_  
Verónica Dahl  
Senior Supervisor

\_\_\_\_\_  
Edward Stabler U of Western Ontario  
External Examiner

\_\_\_\_\_  
Michel Boyer, Université de Montréal  
Supervisory Committee Member

\_\_\_\_\_  
Robert Hadley  
Supervisory Committee Member

\_\_\_\_\_  
Richard DeArmond. DLLL

\_\_\_\_\_  
August 12, 1987  
Date of Approval

PARTIAL COPYRIGHT LICENSE

I hereby grant to Simon Fraser University the right to lend my thesis, project or extended essay (the title of which is shown below) to users of the Simon Fraser University Library, and to make partial or single copies only for such users or in response to a request from the library of any other university, or other educational institution, on its own behalf or for one of its users. I further agree that permission for multiple copying of this work for scholarly purposes may be granted by me or the Dean of Graduate Studies. It is understood that copying or publication of this work for financial gain shall not be allowed without my written permission.

Title of Thesis/~~Project/Extended Essay~~

GENERATING SPANISH CLITICS with  
CONSTRAINED DISCONTINUOUS GRAMMARS

---

---

Author:

(signature)

CHARLES G. BROWN

(name)

1987 Aug 27

(date)

## Abstract

A synthesis of Discontinuous Grammars (DGs) and Government and Binding (GB) theory is presented. DGs are a logic grammar formalism with the power to express free word order in rewriting rules, and the power to skip over, leaving unanalysed, elements in the analysis of sentences. They have both the declarative interpretation of rewriting rules, and a procedural interpretation due to the automatic theorem prover built into Prolog. GB theory has great explanatory power for the expression of linguistic phenomena. The transformational component of GB theory describes the movement of elements in sentences. The components and principles of GB theory provide a principled explanation of the learnability of language.

A grammar for the generation of tensed Spanish sentences with object clitics is presented. The grammar demonstrates the synthesis in a single logic grammar of a Base Component and a Transformational Component conforming to the precepts of GB theory. Some of the sentences generated, the dative clitic constructions called benefactive datives, can be considered idiomatic expressions. Semantic representations for these constructions were developed and shown to share characteristics with dative of interest and ethical dative constructions.

The problem of the syntax-semantics interface was solved through the introduction of generalised pre-lexical processing of the Environment of Syntax for each maximal projection. This effectively allows the syntax of generation to be independent of the semantic representation of sentences. The decomposition of the semantic representation is not

referred to in the grammar rules, a departure from previous practice in logic grammars that support semantic interpretation. The addition of pre-lexical processing also allows the grammar to generate both active and passive versions of a sentence where the lexicon allows. This processing is quite general and may easily be applied to generation grammars for other languages.

The original formulation of Discontinuous Grammars did not licence generation. A new interpretation of Discontinuous Grammar rules is introduced that both allows generation and has linguistic motivation based on GB theory. Constraints on DGs were extended with a new kind of active constraint which replace the passive filters of GB theory. A compiler for Discontinuous Grammar rules under the new interpretation is presented.

## Table of Contents

<b>Approval</b>	ii
<b>Abstract</b>	iii
<b>Table of Contents</b>	v
<b>1. Introduction: Natural Language Generation, Logic Grammars, and Clitics</b>	1
1.1. Government and Binding Theory. The linguistic framework	4
1.1.1. Automatic Generation of Natural Language	5
1.1.2. The Semantic Representation	7
1.1.3. Syntax and Semantics	8
1.1.4. Constrained Discontinuous Grammar Rules	9
1.2. Logic Grammars	12
1.3. Issues about Clitics	21
1.3.1. Clitics in Spanish	22
1.3.2. Clitics and Empty Categories	24
1.3.3. Base generation of clitics	27
1.4. Summary	28
<b>2. The Semantics of Spanish Clitics</b>	30
2.1. Clitic Chains and the OVERT feature	31
2.2. Direct Object Clitics	33
2.2.1. Conceptual graph representation	34
2.2.2. Emphatic and nonemphatic pronominal NPs	36
2.2.3. Dialectal direct object doubling	38
2.2.4. Topicalised direct objects	39
2.3. Impersonal-se Constructions	41
2.4. Indirect Object Clitics	42
2.4.1. Dative clitic doubling	43
2.4.2. Possessive dative	44
2.4.3. Dative of interest	47
2.4.4. Ethical dative	50
2.4.5. Benefactive dative	52
2.5. Summary	54
<b>3. Syntax, Semantics, and the Environment of Syntax</b>	56
3.1. Introduction	56
3.2. The role of X-Bar Theory	58
3.3. The Transformational Component of GB Theory	60

3.4	Topicalisation - an example of a transformation	63
3.5	Transformations and semantics	64
3.6	A model of environment of syntax for NL generation	66
3.6.1.	The components of context	67
3.6.2.	Consulting context from NP	70
3.6.3.	Consulting context from VP	73
3.6.4.	Consulting context from PP	75
3.6.5.	Consulting context from CP	75
3.6.6.	Consulting context from IP	76
3.7	Related work in Automatic NL Generation	77
3.8	Towards a more powerful automatic NL generation system	82
<b>4.</b>	<b>Clitics as Inflections</b>	<b>85</b>
4.1	Introduction	85
4.2	The Nature of Clitics	85
4.2.1.	Separable and Inseparable Affixes	86
4.2.2.	Proximity to the verb	86
4.2.3.	Absolute order of clitics	88
4.2.4.	Clitics are not NPs	89
4.2.5.	Accents and inflections	91
4.2.6.	Explanatory adequacy	92
4.2.7.	Possibilities for clitics	92
4.3	Base Generation of Clitics	92
4.3.1.	Object deletion - Rivas	93
4.3.2.	Clitic chains - Hurtado	93
4.3.3.	Clitic generation in VP	94
4.3.4.	Subject clitic generation in AGR	94
4.4	The Inflection Hypothesis	95
4.4.1.	Paradigmatic conjugation	95
4.4.2.	Base generation of clitics in INFL	96
4.4.3.	Transformations involving INFL	98
4.4.4.	Imperatives	99
4.4.5.	Infinitives	101
4.4.6.	Gerunds	102
4.4.7.	Past participles	102
4.4.8.	Clitic promotion	102
4.5.	Some Observations	103
<b>5.</b>	<b>Generating D-structure for Spanish</b>	<b>104</b>
5.1	X-Bar Theory	105
5.2.	Case and Theta Theory in D-structure	109
5.3	Getting Features from the EOS	109
5.3.1.	Informants at the CP level	110
5.3.2.	Informants at the IP level	111
5.3.3.	Informants at the VP level	112

5.3.4. Informants at the NP level	113
5.4. Generating Impersonal-se D-structure	114
5.5. Generating Benefactive Dative D-structure	116
5.6. Generating True Passive D-structure	121
5.7. Non-deterministic Generation	123
5.8. Comments	124
<b>6. A New Formulation for Discontinuous Grammars</b>	<b>126</b>
6.1. Discontinuous Grammars	126
6.2. A New Formulation for Discontinuous Grammars	137
6.3. An Implementation of a Discontinuous Grammar Compiler under the New Formulation	139
6.3.1 Adding the derivational context	140
6.3.2 Simultaneous application of rules	142
6.4 Summary	145
<b>7. Transformations and Spanish S-structure</b>	<b>146</b>
7.1. Move XP Transformations	146
7.1.1 A Move-NP transformation	147
7.2 Move-X Transformations	150
7.2.1 The Verb to INFL transformation	150
7.3 Minor Movement Rules	152
7.3.1 The Affix Hopping transformation	152
7.3.2 Clitic Hopping	154
7.4 Clitic Promotion	156
7.5 Overlapping and Nested Movement	157
7.6 Computing Transformations	159
<b>8. Remarks and Conclusions</b>	<b>160</b>
8.1. Some Remarks on Automatic NL Generation	161
8.2. Conclusions and Contributions	163
8.2.1 Interpretation of DGs	163
8.2.2. Implementation of a CDG Compiler	164
8.2.3. The Environment of Syntax	165
8.2.4 A Working Transformational DG	166
8.2.5 Dynamic Constraints on NL DGs	167
8.2.6 Generation of Passives and Actives	168
8.2.7 Non Deterministic Generation	169
8.2.8 Base Generation of Clitics	169
8.2.9 Analysis of Clitics as Inflections	170
8.2.10 Representation of Special Clitic Constructions	171
8.2.11. Pre-Lexical Processing	172
8.3. Remarks on Some Open Questions	173
8.3.1. The Theta Criterion and Chain Composition	173
8.3.2. Temporal Relations and Mood, Tense, Voice, and Aspect	174
8.3.3. Expression of Adjunction in Logic Grammar Rules	175



8.3.4. Clitic Promotion	176
8.3.5. Projection	176
8.4. Final Comments	177
<b>Appendix A. A Glossary of Some Concepts of GB Theory</b>	<b>178</b>
<b>Appendix B. A Compiler for Constrained Discontinuous Grammar Rules</b>	<b>183</b>
<b>Appendix C. Arguments that Clitics are not NPs</b>	<b>188</b>
C.1. Morphological differences	188
C.2. NP forms do not distinguish Case	189
C.3. NPs take stress	189
C.4. Phonological considerations	190
C.5. Surface order of NPS	190
C.6. NPs take modifiers	191
C.7. NPs may be conjoined	191
C.8. Transformations apply only to NPs	191
C.9. Only clitics may double NPs	192
<b>References</b>	<b>194</b>

1. Since the defense of this dissertation, it has become apparent that a new and more appropriate name for the formalism and grammar described is **Static Discontinuity Grammar**. The class of grammars described by the formalism in Chapter 6 is distinct from the class of grammars described in *Constrained Discontinuous Grammars - a linguistically motivated tool for processing language*, ([Dahl and Saint-Dizier 86] ). The Constrained Discontinuous Grammars described in [Dahl and Saint-Dizier 86] simply augment the general Discontinuous Grammar formalism with constraints. A Static Discontinuity Grammar has the special feature that skips do not move, that is, they are *static*. The constraints described in the dissertation exploit these static skips (as described in Chapters 6 and 7). The first description of this class of grammars may be found in the article by Verónica Dahl: *Gramáticas Discontinuas - una herramienta computacional con aplicaciones en la teoría de rección y ligamiento* in the journal *Revista Argentina de Lingüística*, volume 2 number 2, 1986, although at that time the term Static Discontinuity Grammar was not yet coined.

Accordingly, by agreement with all the authors involved, the term **Static Discontinuity Grammar** should replace those occurrences of **Constrained Discontinuous Grammar** that do not specifically refer to the Dahl and Saint-Dizier report.

2. The following addenda should be inserted where indicated in the bound copy of the dissertation.

page 56 The following footnote should be added to Chapter 3:

*The forerunner of the work described in this chapter was the development, by the Aymaras research group at Simon Fraser University, of the processgraph algorithm and program described in [Brown, Dahl, et al 1986] and the exploitation of the Stowell's version of X-bar Theory, a lexicon with thematic-role information, and Sowa's conceptual graphs as a semantic representation. In work supported in part by IBM Canada Ltd, we produced a demonstration system [Brown, Pattabhiraman, et al 86] [Brown, Dahl, et al 86]. This chapter extends that work and introduces the general mechanism of pre-lexical processing and the concept of environment of syntax.*

page 170 Add the following under example (8.5):

In addition, the following sentences with overt subject pronominalisation could be generated in unconstrained generation. Such pronominalisation might be made to clarify that the subject is male, and not female. As such it is a context-dependent pronominalisation. Only trivial changes to the generation grammar are needed to permit or include this kind of pronominalisation.

- (8.5.1.a) El llamó a la niña. (Real Academy Spanish).  
*He called the child.*
- (8.5.1.b) El la llamó a la niña. (clitic doubling, dialectal/emphatic).  
*He called the child.*
- (8.5.1.c) El la llamó. (object pronominalisation).  
*He called her.*
- (8.5.1.d) El la llamó a ella. (emphatic object pronominalisation).  
*He called HER.*

page 193 Add the additional acknowledgement:

*The origins of this thesis go back to the work done by the Aymaras research group under Verónica Dahl in a Shared University Research project funded by IBM Canada Ltd. I gratefully acknowledge the co-authorship of T. (Pat) Pattabhiraman in designing the processgraph algorithm for consulting a conceptual graph for the demonstration program (based on a fragment of English) of that project.*

*Pat designed a datastructure for conceptual graphs and then implemented our algorithm based on that datastructure. In addition he and Pierre Massicotie continued to improve the implementation of the algorithm, which we carried through many iterations for the Aymaras project. It is a tribute to the robustness of Pattabhiraman's datastructure and to the implementation of the algorithm by Pat and Pierre that I was able to carry it over unchanged to my thesis work on Spanish.*

*Pierre developed a complete and detailed (ie usable) transformational syntax of the English auxiliary system for the Aymaras project. I was able to adapt that system easily to my treatment of Spanish.*

Errata October 1, 1988

The following corrections should be made to the bound edition of the dissertation.

page 70 In section 3.6.2, the reference in the second paragraph (which begins *These tasks ...* ) should be to [Brown, Dahl, et al 1986].

page 120 In diagram 5.20 the verb escapar should be perder.

page 170 The example 8.5 should be replaced in its entirety with the following:

(8.5.a) Juan llamó a la niña. (Real Academy Spanish).

*Juan called the child.*

(8.5.b) Juan la llamó a la niña. (clitic doubling, dialectal/emphatic).

*Juan called the child.*

(8.5.c) Juan la llamó. (object pronominalisation).

*Juan called her.*

(8.5.d) Juan la llamó a ella. (emphatic object pronominalisation).

*Juan called HER.*

(8.5.e) Llamó a la niña. (null-subject pronominalisation.)

*He called the child.*

(8.5.f) La llamó a ella. (null-subj pronom with emphasis.)

*He called HER.*

page 197 The reference [Dahl and Brown 86] should have as authors: Verónica Dahl, Charles Brown, and Sharon Hamilton.

page 197 The reference [Dahl and Saint-Dizier 86] should have as its full title: *Constrained Discontinuous Grammars - a linguistically motivated tool for processing language*. This reference is also available as INRIA Report no. 573, 1986

## Chapter 1

### Introduction: Natural Language Generation, Logic Grammars, and Clitics

This thesis presents a synthesis of Discontinuous Grammars (DG) [Dahl and Abramson 84] [Dahl 84] and Government and Binding (GB) theory [Chomsky 82a] [Chomsky 86]. DGs are a logic grammar formalism and have great expressive power. The power, for example, to express free word order in rewriting rules, and the power to skip over leaving unanalysed elements in the analysis of sentences. They also have both the declarative interpretation common to rewriting rules and a procedural interpretation due to the automatic theorem prover built into the Prolog language. GB Theory has great explanatory<sup>1</sup> power for the expression of linguistic phenomena. The transformational component of GB Theory describes the movement of elements in sentences. The components and principles of GB Theory describe natural language (NL) generation very succinctly and in doing so, provide a principled explanation of the learnability of language.

One goal was then, to bring together these two formalisms. The one (DG) dealing with a grammar formalism and logic programming. The other (GB) dealing with natural language generation. DGs appeared to have the potential for expressing the

---

<sup>1</sup>Explanatory power has been described in *Lectures on Government and Binding* [Chomsky 82a] as the ability of a grammar to account for the fact that "Knowledge of language is acquired on the basis of the evidence (i.e. the data or utterances) available. A grammar that has explanatory adequacy must meet the three criteria proposed by Chomsky given in Appendix A.

transformations of GB Theory. In particular, the ability to express nested<sup>2</sup> and overlapping transformations not possible<sup>3</sup> with other logic grammar formalisms

Another goal was to develop a generation logic grammar that could generate tensed Spanish sentences with object clitics. I believe that this is the first time that such constructions have been generated with a logic grammar. Some of these, the dative clitic constructions called benefactive datives, can be considered idiomatic expressions. Semantic representations for these constructions were developed and shown to share characteristics with dative of interest and ethical dative constructions. The synthesis in a single logic grammar for NL of a Base Component and a Transformational Component conforming to the precepts of GB Theory is demonstrated in my generation grammar for Spanish

Two major problem areas existed. One was the syntax semantics interface in the logic grammar. This was solved through the introduction of generalised pre-lexical processing of the Environment of Syntax (or EOS discussed in Chapter 3) for each maximal projection<sup>4</sup>. This effectively allows the syntax of generation to be independent of the semantic representation of sentences. In my grammar, the decomposition of the semantic representation is not referred to in the rules. This is a departure from previous practice in logic grammars for NL that support semantic interpretation. The addition of prelexical processing also allows the grammar to generate both active and passive versions of a

---

<sup>2</sup>Nested transformations are those in which the scope of one transformation is entirely within the scope of another transformation

<sup>3</sup>In a recent presentation but not in the published paper [Joshi 87] discussed the need to describe nested and overlapping dependencies. The framework was his own work on Tree Adjoining Grammars to express such dependencies within the linguistic framework of Generalised Phrase Structure Grammars

<sup>4</sup>Maximal projection is a linguistic term for the major categorial components (eg. NP VP) of the syntax of a natural language. More details are given in Chapter 2.

sentence where the lexicon allows. This processing is quite general and may easily be applied to generation grammars for other languages.

The other problem was to find an interpretation of DGs with linguistic motivation that allowed mapping of DG rules to Prolog clauses. The DG rules were wanted to express movement transformations of GB Theory (Phrase structure rules could be expressed with Prolog clauses, without the need for DG rules.) The linguistic theory, that is GB Theory, pointed the way to this new interpretation. To this end, Constrained Discontinuous Grammars (CDG) [Dahl and St-Dizier 86] were extended in [Dahl Brown et al 86] to incorporate the new linguistic interpretation of DGs, and extended with a new kind of active constraint on NL generation. The active constraints on generation replace the passive filters of GB Theory and in this respect the grammar departs from GB Theory. A related problem for the new interpretation of DGs was the inability of DGs as originally formulated to generate sentences, although they worked well for analysis. This problem was solved as well with the new interpretation of DG rules. A compiler for DG rules under the new interpretation is introduced in Chapter 6.

The work in this thesis combines the two disciplines of Computing Science and Linguistics. Because of this, readers may wish to consult Appendix A. Brief explanations of terms from Linguistics and Computing are given along with references to further background reading.

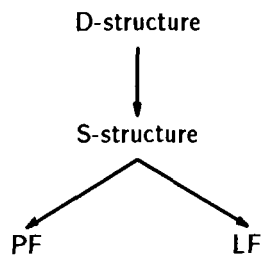
The remainder of this Chapter discusses GB Theory and automatic NL generation, logic grammar formalisms, and linguistic issues of Spanish clitic constructions. Chapter 2 examines the semantics of some Spanish clitic constructions. Chapter 3 introduces the EOS and my ideas on prelexical processing. Chapter 4 presents my theory of Spanish

clitics as verbal inflections Chapter 5 discusses the logic grammar rewrite rules and prelexical processes for the clitic constructions of my generation logic grammar. Chapter 6 examines DGs and the motivation for a new interpretation of DG and CDG rules, and presents a compiler based on the new interpretation. Chapter 7 discusses the kinds of transformations used in the generation grammar and presents CDG rules for these transformations Chapter 8 summarises the contributions of this thesis and discusses some areas for future research.

## 1.1. Government and Binding Theory: The linguistic framework

Government and Binding (GB) theory has a Base component that generates D-structure (or deep syntactic structure) and a Transformational component that restructures D-structure into S-structure. Two further components act on S-structure to produce Phonological Form (PF) and Logical Form (LF).

(1.1)

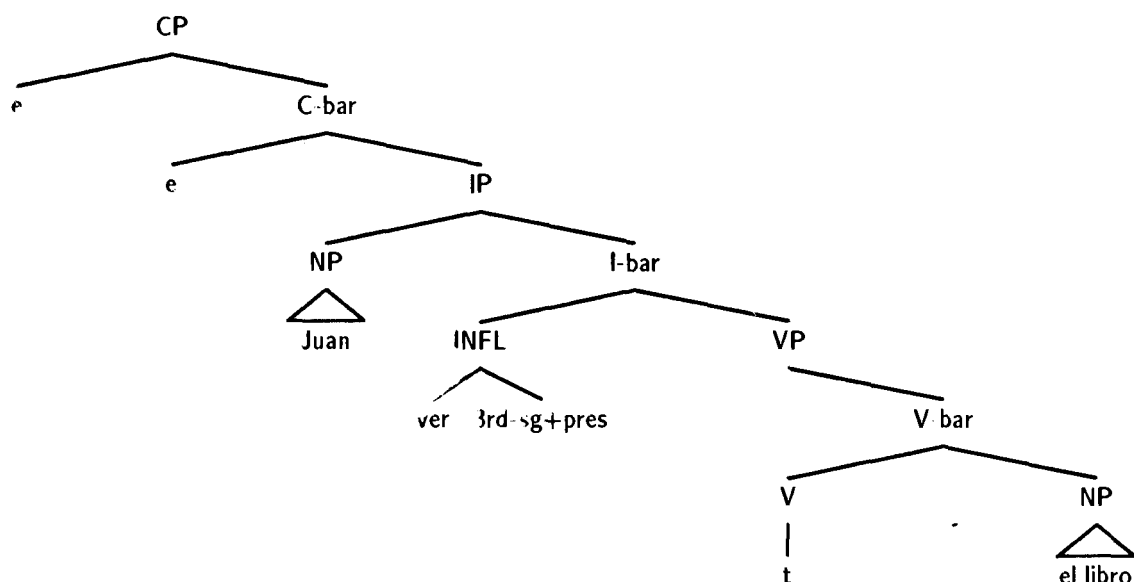


The Base component consists of the categorial rules of the grammar and the lexicon. The Transformational component provides for movement of elements from one position in the syntactic structure to another.

The scope of the logic grammar is to generate S-structure such as (1.2) (and not to

generate the final morphological or orthographic forms.) Broad linguistic coverage is provided, including active and passive sentences, NP-movement, affix-hopping, verb-to-INFL movement, and constructions specific to Spanish such as benefactive datives, impersonal-se subjects, and clitic doubling.

- (1.2) Juan ver AGR+pres el libro. (S-structure)  
 Juan sees AGR+pres the book.

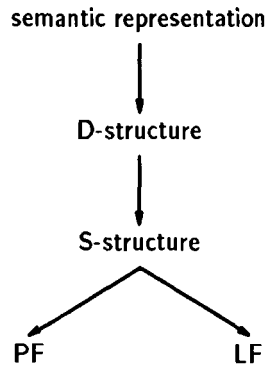


### 1.1.1. Automatic Generation of Natural Language

Generation is assumed to proceed from a predetermined semantic representation of the meaning of a single sentence. Introduction of this abstract level of semantic representation implies an order of generation from semantics to syntax. Generative Semantics adopted such an approach in the 1960's and early 1970's. However, that approach posited transformations on the semantic structure, intermixed with lexical substitutions based on lexical decomposition to arrive at a final syntactic surface structure. No deep structure of syntax was admitted to exist in Generative Semantics. That is not the case here. The semantic representation will be consulted by the grammar in order to generate D-structure. This is the subject of Chapter 3.



## (1.3)



Automatic generation of sentences might form one part of a system for the automatic translation of text between languages. The model of translation in this case has an intermediate semantic representation (sometimes called an interlingua)<sup>5</sup> The interpretive semantics of [May 84] (within GB Theory) is not appropriate to this model of translation because the *logical form* (or LF) is derived from S-structure and seems to be language specific

Another proposed use of an automatic NL generation system is to generate operating system messages in several languages from a common semantic representation. Similarly, a question answering system could make use of an automatic NL generation system to generate responses to queries from an internal representation of the responses

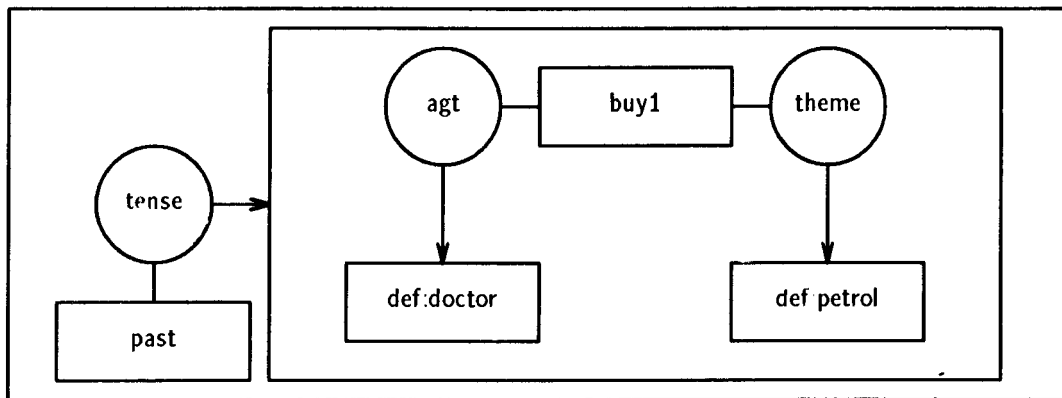
---

<sup>5</sup>As opposed to the transfer model, which makes no use of semantic representation.

### 1.1.2. The Semantic Representation

The semantic representation chosen for this work has been the Conceptual Graph of [Sowa 84] . but it is felt that the grammar is effectively independent of the semantic representation. This is demonstrated in Chapter 3 where a theory of the semantic-syntax interface called the Environment of Syntax is proposed. An example of a conceptual graph and the sentence it represents is given in (1.4).

- (1.4) **El médico compró la bencina.**  
*The doctor bought the petrol.*



**A(z) the(x) the(y) [doctor(x) and petrol(y) and buy1(z)  
 and agent(z,x) and theme(z,y)]**

Conceptual graphs are a graphical representation of predicate logic. They are composed of concepts and relations. Concepts are meant to represent neural precepts. Relations specify the roles that the precepts play. More formally, conceptual graphs are finite, connected, bipartite graphs. The two kinds of nodes are concepts (boxes) and conceptual relations (circles). Every conceptual relation has one or more arcs which must be linked to some concept. A conceptual graph may be embedded as a concept in another conceptual graph. Each conceptual graph asserts a proposition (and a single concept may itself be a conceptual graph). The proposition may be shown graphically, or mapped to a first order

logic<sup>6</sup>.

Using (1.4) as an example<sup>7</sup>, the concepts in the embedded proposition are **buy1**, **def:doctor**, and **def:petrol**. The concept **past** is in a higher proposition. Relations are **agt** and **theme** in the embedded proposition, and **tense** in the higher proposition. The embedded proposition may be mapped to the logical proposition shown below the graph.

Looking at such propositions linguistically, we can see that the concept **buy1** in (1.4) may be thought of as a semantic predication with arguments **def:doctor** and **def:petrol**. However the arguments are in named positions in the graphs, the names being given by the relations **agent** and **theme** respectively. These named argument positions will be useful to us. As we shall see later, the conceptual relation names may correspond to  $\Theta$ -roles<sup>8</sup> in some cases providing a link with the subcategorisation features of the lexicon.

### 1.1.3. Syntax and Semantics

A basic tenet of GB Theory is the principle of Autonomous Syntax. This holds that the rules governing the structure of sentences (the syntax) in a language are independent of the meaning of the sentence.

I have assumed an independent representation of the meaning of a sentence. For

---

<sup>6</sup>The mapping from CG to logic expression is assumed by [Sowa 84] to exist in the form of the operators (see pages 104-105 of Sowa).

<sup>7</sup>English is used as a semantic meta-language in all the semantic representations in this thesis.

<sup>8</sup>Thematic roles, or  $\Theta$ -roles, are the semantic roles played by the arguments of a sentential proposition. They correspond to Fillmore's Cases Relations [Fillmore 68], to Gruber's Thematic Relations [Gruber 65], or to Jackendoff's Thematic Relations [Jackendoff 72] [Jackendoff 83]. Typical  $\Theta$ -roles are **agent-of-action**, **goal-of-action**, **patient-of-action**, **destination**, **recipient**, etc.

automatic generation of NL the grammar of the language must draw on this semantic representation. However, as I see it more than the semantic representation is involved here. Also involved are the semantic types of concepts, a semantic-type hierarchy, the semantic aspects of the lexicon, and pragmatics of the context of discourse. The semantic representation may not be merely restructured into syntactic structure. (This was amply demonstrated during the Generative Semantics era.)

I have postulated the EOS as a level of reference that the grammar may draw upon for the generation of a single sentence. Its data structures represent the factors mentioned above. It contains the components listed below. How these components are consulted by the grammar in a regular way will be the subject of Chapter 3.

- An lexicon (a local lexical context).
- A categorial type (a local syntactic context)
- A semantic representation for one sentence (a local semantic context).
- A semantic-type hierarchy (a local pragmatic context).
- An R expression 'stack' (a local discourse and sentential context).

I will propose, in Chapter 3, that each  $XP^9$  generated in D-structure requires one consultation with the EOS, as illustrated graphically in (15)

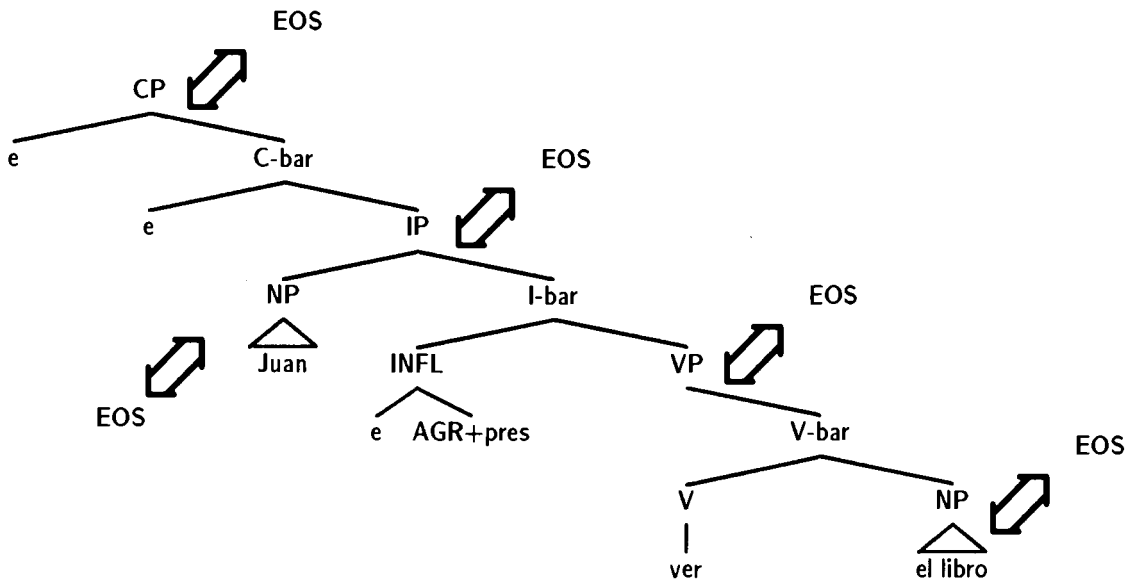
#### 1.1.4. Constrained Discontinuous Grammar Rules

The Transformational component of GB Theory consists of the general rule, Move  $\alpha$ , which allows categorial components of the syntax to move from one position in the abstract syntactic structure to another. I have used CDG rules to express such transformations in the logic grammar.

---

<sup>9</sup>XP is the general name for maximal projection. One of NP, VP, PP, etc.

(1.5) Juan AGR+pres ver el libro. (D-structure)



CDG rules are rules of the general form:

**a, skip, b, skip, . . . , c --> a', skip, b', skip, . . . , c'**

The interpretation of these rules is the simultaneous application of the rules:

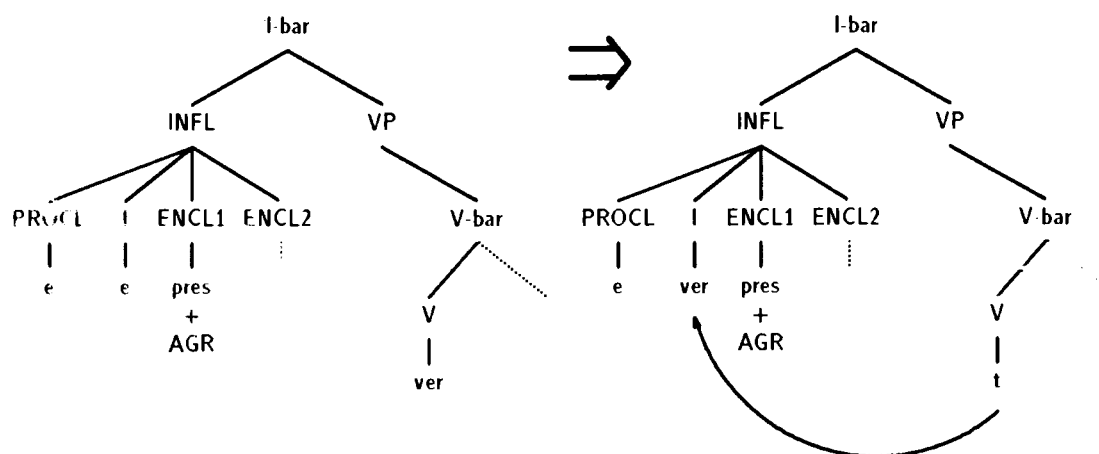
**a --> a'**  
**b --> b'**  
 . . .  
**c --> c'**

all of which must succeed if the CDG rule is to succeed. In addition, CDG rules can both provide control over which transformations take place through active filters, and also provide the derivational tree structure of the syntax upon which constraints (such as government of empty categories) apply. Control and constraints in CDGs will be discussed in Chapter 7.

In order to generate Spanish sentences with clitics, CDGs have been used for several

transformations. For example<sup>10</sup>, consider the transformation<sup>11</sup> shown in (1.6) that moves a verb (an  $X^0$  category) from its D-structure position in VP to the position of the inflected verb in INFL (another  $X^0$  position). A trace ( $t$ ) is left in the position vacated by the moved verb.

(1.6)



In order to express this transformation with a CDG rule, the rule that normally rewrites the V position in INFL to e (empty):

(1.7)  $x(\text{infl}, \text{PT}, \text{G}, \_) \rightarrow [e], \{\text{PT} = \dots [i, e]\}.$

is replaced with the CDG rule:

(1.8)  $x(\text{infl}, \text{PT1}, \text{G}, \_), \text{skip}, x(\text{vp}, \text{PT2}, \text{G}, \text{W}) \rightarrow x(\text{vp}, \text{PT1}, \text{G}, \text{W}), \text{skip}, x(\text{vtrace}, \text{PT2}, \text{G}, \_).$

This expresses the simultaneous application of two rules:

<sup>10</sup>For the reader not familiar with Prolog, the notation used in this example is explained in Chapter 6.

<sup>11</sup>This transformation has been proposed in Barriers [Chomsky 86]. I have assumed this analysis of verbal affixes of tense and agreement rather than the earlier affix-hopping analysis.

- (1.9.a)  $x(\text{infl}, \text{PT1}, \text{G}, \_) \rightarrow x(\text{vp}, \text{PT1}, \text{G}, \text{W}).$   
 (1.9.b)  $x(\text{vp}, \text{PT2}, \text{G}, \text{W}) \rightarrow x(\text{vtrace}, \text{PT2}, \text{G}, \_).$

Note that the rules (1.9) share variables through unification. Rule (1.9.a) will be followed by the application of the rule:

- (1.10)  $x(\text{vp}, \text{PT1}, \text{G}, \text{W}) \rightarrow [\text{W}], \{\text{PT1} = \_[\text{v}, \text{W}]\}.$

which gives the lexical item,  $[\text{W}]$ , and rule (1.9.b) leads to the application of the rule:

- (1.11)  $x(\text{vtrace}, \text{PT2}, \text{G}, \_) \rightarrow [\text{t}], \{\text{PT2} = \_[\text{v}, \text{t}]\}.$

which results in a terminal trace,  $[\text{t}]$ .

All the transformations used in the generation grammar are discussed in more detail in Chapter 7

## 1.2. Logic Grammars

Logic based grammars are tools for describing powerful rewriting rules. They are, as the name implies, based upon logic programming. Colmerauer [78] developed the first logic grammar formalism in 1975. His *Metamorphosis Grammars* (MG) formalism allowed *type-zero like* rewrite rules with more than one symbol on the LHS to be written directly in a way that hides string manipulation from the user.

Logic grammars differ from type-zero grammars in the following important ways [Dahl 85]:

- The form of the grammar symbols may include arguments.
- Variables may be used. They are bound to values through unification.
- Rules may include tests (predicates on database and linguistic information).
- The Prolog theorem-prover endows the rules with a procedural semantics. The procedural interpretations allow the grammars to become parsers and synthesisers.

A Metamorphosis Grammar,  $G$ , is defined<sup>12</sup> as a 4-tuple  $(V_N, V_T, R, P)$  where<sup>13</sup>.

$V_N$  the set of nonterminal symbols,  $V_N \subseteq \hat{H}[F]$   
 $V_T$  the set of terminal symbols,  $V_T \subseteq H[F]$   
 $V = V_N \cup V_T$   
 $R$  the set of starting symbols,  $R \subseteq V_N$   
 $P$  the set of productions of the form:

$$\alpha_1, \alpha_2, \dots, \alpha_m \rightarrow \beta_1, \beta_2, \dots, \beta_n$$

with:  $m \geq 1, n \geq 1$ .

$$\alpha_i \in V \text{ for } 1 \leq i \leq m,$$

$$\beta_j \in V \text{ for } 1 \leq j \leq n.$$

The language  $L(G)$  associated with  $G$  is<sup>14</sup>:

$$L(G) = \{\omega \in V_T^* \mid s \rightarrow^* \omega \text{ for } s \in R\}$$

The implementation of MG permitted only *normalised* productions. Normalised MG productions must start with a non-terminal symbol. There is no loss of generality in this restriction [Colmerauer 78]. Some example MG productions appear in (1.12).

Since the development of logic based meta-grammars by [Colmerauer 78] (MGs) and the development by [Dahl 77] [Dahl and Sambuc 76] of techniques for using them, research

<sup>12</sup> [Colmerauer 78]

<sup>13</sup>The symbols of a logic grammar, *logic terms*, are *functors* with zero or more arguments. The set of logic terms  $F$  is a denumerable set of functors of zero or more arguments.  $\hat{H}[F]$  is the set of logic terms that can be constructed from  $F$ .  $H[F]$ , the **Herbrand Universe**, is the set of logic terms without variables.

<sup>14</sup>

Here:

$$V_T^* = \bigcup_{i=0}^{\infty} V_T^i$$

which is the Kleene closure of the set  $V_T$ , and

$$\rightarrow^* = \text{the reflexive, transitive closure of } \rightarrow.$$



(1.12)

 $a, [b] \rightarrow [b], a.$  $\text{noun-root}(R), \text{plural-mark} \rightarrow [W], \text{concatenate}([X], [s], W).$  $\text{relative-marker}, \text{noun-phrase}, \text{moved-dobj} \rightarrow \text{rel-pronoun}, \text{noun-phrase}.$ 

has proceeded in three areas: **firstly**, formalisms for describing grammars (representing language) [F. Pereira 81] [Dahl and Abramson 84] [Dahl 84]; **secondly**, analysis of Natural Language(NL) to provide an internal representation [Dahl 79] [McCord 82] [Coelho 79] [Dahl and McCord 83] [F. Pereira and Warren 80] [Pique 82] [Saint-Dizier 86b] [Saint-Dizier 86a]; and **thirdly** the development of applications techniques for specific linguistic problems [Dahl 79] [Dahl and McCord 83] [McCord 84] [Dahl and Abramson 84].

The first area of research, formalisms for describing grammars, has seen proposals by, for example, [F. Pereira and Warren 80] (DCGs), [F. Pereira 81] (XGs - Extraposition Grammars), [Dahl 84], and [Dahl and Abramson 84] (DG - Discontinuous Grammars), [Sabatier 84] (ZGs - Puzzle grammars), and [Stabler 87] (RLGs - Restricted Logic grammars)

F Pereira and Warren [80] developed a restricted form *Definite Clause Grammars* (DCG) of Colmerauer's formalism motivated by ease of implementation. Some expressive power is lost in DCGs but the ability to describe type-zero grammars is retained.

The DCG formalism only permits productions with a single nonterminal on the left:

$$\alpha \rightarrow \beta_1 \beta_2 \dots \beta_n$$

$$\text{with: } \alpha \in V_N, n \geq 0, \\ \beta_i \in V \text{ for } 1 \leq i \leq n$$

Two examples of DCG productions are given in (1.13).

(1.13)

**n-phrase(np(D,N,Rel)) --> determiner(D), noun(N), rel-clause(Rel).**  
**determiner(det(D)) --> [D], {is-determiner(D)}.**

Each production is rewritten as a Horn (Prolog) clause automatically in such a way that each non-terminal gains two new arguments corresponding to input and output lists. The Prolog clauses corresponding to (1.13) are given in (1.14). Conditions (procedure calls) may be placed on the production in a straightforward manner by placing the predicate name in curly brackets as in the second production rule in (1.13). Note that the procedure call **is-determiner** does not gain the two extra arguments, as it does not consume an input element.

(1.14)

**n-phrase(np(D,N,Rel),\_1,\_4) :- determiner(D,\_1,\_2),**  
**noun(N,\_2,\_3),**  
**rel-clause(rel,\_3,\_4).**  
**determiner(det(D,\_1,\_2)) :- det(D,\_1,\_2),**  
**is-determiner(D).**

More expressive power has been provided by Pereira's *Extrapolation Grammars* (XG). It appears that efficiency may be a problem with this formalism. XGs allow the use of *gaps*<sup>15</sup> on the LHS of a production. The gaps are rewritten *in the order in which they occur* on the *rightmost end* of the RHS of the production. Gaps refer to unspecified strings of symbols in the production. This makes possible the description of left-extrapolated constituents. There are however restrictions on the power of XGs to describe left extrapolation. The form of XG productions is given in (1.15) below. The different gaps in XGs may be represented through a symbol, *gap(X)*, where the X denotes the gap

---

<sup>15</sup>The term *gap* is used here in the sense of *unanalysed material*. This is not the same as the linguistic use of the term as an empty category or syntactic position, as in parasitic gap.

in question<sup>16</sup>.

(1.15)

$$\alpha_1.\text{gap}(X_1).\alpha_2.\text{gap}(X_2).\dots.\text{gap}(X_{n-1}).\alpha_n \rightarrow \beta.\text{gap}(X_1).\text{gap}(X_2).\dots.\text{gap}(X_{n-1}).$$

where:

$\alpha_1$  must be a non-terminal.

More recently, the RLG (Restricted Logic grammar) of [Stabler 87] has shown how the rules for moved constituents of XGs can be augmented with *switch rules* similar to the rules of the Marcus parser [Marcus 80]. More importantly, Stabler introduces the use of the c-command form of Government (of [Chomsky 82a]) and the use of Subjacency to constrain both leftward and rightward movement of constituents in RLGs.

The Discontinuous Grammar<sup>17</sup> (DG) formalism first conceived and implemented by Dahl, was later applied to coordination by Dahl and McCord only to be dropped in favour of other work. Dahl and McCord [83a] and Dahl and Abramson [84] developed an alternative implementation, and the formalism's uses were studied by [Dahl 84]. DGs have more generality than XGs and have been applied to a NL analysis system to provide top-down or bottom-up syntactic parsing [Popowich 85].

Both DGs and XGs can be described as 5-tuples  $(V_N, V_T, \Gamma, \Sigma, P)$  where  $\Gamma$  is the set of gap symbols, and  $\Gamma \cap V = \phi$ .

---

<sup>16</sup>The notation used here is different from that used by Pereira. In Pereira's notation, the gaps are denoted by the three dots functor. No gaps appear on the RHS of the production, but they are implicitly rewritten in the order in which they occur on the rightmost end of the production.

$$\alpha_1. \dots .\alpha_2. \dots .\text{etc.} \dots .\alpha_n \rightarrow \beta$$

where:

$\alpha_1$  must be a non-terminal.

<sup>17</sup>Discontinuous Grammars were originally named Gapping Grammars. This term led to some confusion with the term *gap* in linguistic theory, used to denote an empty category (as in *parasitic gap*).

The general form of DG productions is:

(1.16)

$$\alpha_1, \gamma_1, \alpha_2, \gamma_2, \dots, \gamma_{m-1}, \alpha_m \rightarrow \beta_1, \delta_1, \beta_2, \delta_2, \dots, \delta_{n-1}, \beta_n$$

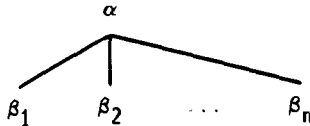
where:

$$\begin{aligned} \alpha_1 &\in V_N \\ \alpha_j, \beta_i &\in V_N \text{ for } 2 \leq j \leq m, 1 \leq i \leq n \\ \gamma_i, \delta_i &\in \Gamma \text{ for } i \leq m-1, j \leq n-1 \end{aligned}$$

DGs allow the possibility of free word order and partially-free word order grammars.

The Puzzle Grammar (ZG) formalism of [Sabatier 84] provides an alternative notation for MGs. In ZGs, *key-trees* of the form shown in (1.17) are used instead of productions.

(1.17)



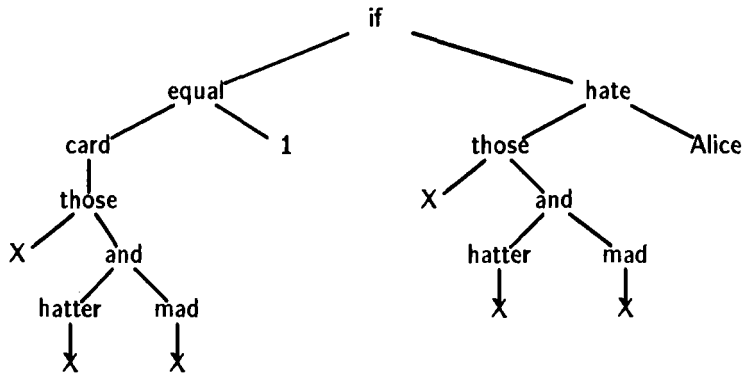
where:

$$\begin{aligned} \alpha &\text{ is a predicate,} \\ \beta_i &\in V_T \cup V_N \cup [ ] \cup \{\text{key-trees}\} \end{aligned}$$

New key-trees are assembled by unifying roots and leaves of key-trees. Constraints may be attached to leaves of key-trees to restrict the assembly. A question of control exists with this formalism. It is not yet clear that efficient application to NL processing is possible.

Now let us turn to the second general area of research in logic grammars, that of providing internal representations for the meanings of utterances. Internal representation was provided by Dahl's formalism for quantification in three-valued logic [Dahl 79] as a functionally compositional semantic representation analysis as shown in the example (1.18).

(1.18) *The mad hatter hates Alice.*



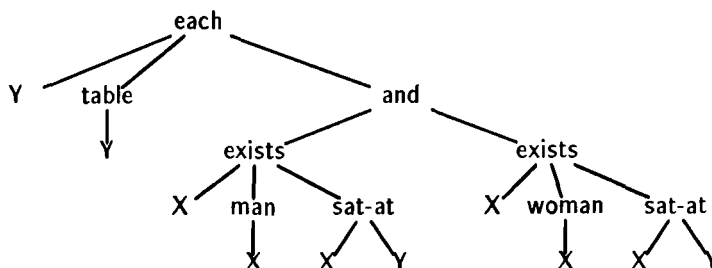
McCord [82] provided a semantic representation by restructuring a syntax tree. This however is a conjunctive form as shown in (1.19).

(1.19) **The man saw John.**

**dcl(man(X) => Y=john & see(X,Y,C) & past(C))**

Dahl and McCord's [83a] metagrammatical system for the treatment of coordination produced both semantic and syntactic structures. Their system included automatic build-up of syntactic and semantic structures. Figure (1.20) gives an example of such a semantic structure, a functionally compositional semantic representation.

(1.20) *A man and a woman sat at each table.*



Abramson [84] isolated the notion of automatic structure buildup in *Definite Clause Translation Grammars*. McCord [84] has recently developed a more complete non-decompositional semantic representation for the EPISTLE project.

Saint-Dizier [86a] has contributed to the semantic representation of adjectives, complex nounphrases, and negation and developed a set of 'tools' for the purpose based on set theory.

Hirschman's [86] Meta Restriction Grammar for NL analysis uses context free rewrite rules, augmented with constraints based on the shape of the partially built parse tree and the input word stream. Automatic build up of the parse tree, basic tree relations, and a set of restriction operator primitives are provided. No semantic representation is included.

Porto and Filgueiras [84] have proposed an intermediate semantic representation determined by rules specific to syntactic categories related to the methods of [Dahl 80] and [Giannesini et al 85].

Lastly we turn to the work on specific linguistic problems that have been addressed with logic grammars. Coordination was treated by [Dahl and McCord 83] in their Modifier Structure Grammars (MSG), and by [Hirschman 86].

Quantifier treatment initially by [Dahl 80], has been extended by [Dahl and McCord 83], and by [Saint-Dizier 86b]. Dahl's formalism for the quantifiers has been adapted to NL database query systems in different languages: French and Spanish by Dahl, Portuguese by [Coelho 79], and English by [F. Pereira and Warren 80]. These adaptations concerned primarily the lexicon.

Dahl's analysis used a three valued logic to deal with presupposition [Dahl 80]. Coelho [79] dealt with presupposition in a highly constrained context using a grammar of discourse. Kaplan [79] dealt with the analysis of presupposition to provide cooperative responses, but not within a logic grammar formalism.

The scope of quantifiers was dealt with by [Dahl 79] in her work on three-valued logic. McCord [82] formalised the ideas of slots and modifiers and developed a more flexible strategy for determining modifier scope. Saint-Dizier developed a subtle treatment of quantifier scoping in [Saint-Dizier 84].

Interpretation of the internal representation, specifically the semantic representations, has been confined to a narrow class of applications, that of database queries. For example, Dahl's three-valued logic system (L3) provides a well-formed set of logic formulas (semantic interpretations of the NL query, *i.e.*, a well-defined set of internal semantic representations), which are interpreted with respect to a given database to provide an answer set or a truth-value [Dahl 80]. Coelho interpreted a similar semantic representations and provided a truth-value, an answer set, or a query to the user based on a tightly defined Grammar-of-Dialogue.

All the aforementioned work on logic-based grammars has focussed on NL analysis. Little work on logic-based NL generation has been done. Boyer and Lapalme [85] have developed a generator for generating from semantic networks and Coelho [79] has investigated the problem of generating from very simple semantic representations. The problem of generating from semantic representations that contain multiple quantifiers has not been investigated. Sharp [85] has developed a strictly syntactically and lexically based translation system using DCG based on the principles of Government and Case Theory including a transformational component.

No work in logic grammars has been done on the treatment of clitics in analysis, semantic representation, or generation. This may be true of all the work in computational linguistics, not only that in logic grammars. Very little work has been reported on pronominal substitution in generation, and again, none that I am aware of in logic grammars. Simmon and Slocum [72] point it out as an unresolved problem in their work on semantic-net based generation. The grammar presented here is the first extensive application of CDGs to date.

### **1.3. Issues about Clitics**

Several issues are important in any discussion of the generation of clitics. After a brief look at the types of Spanish clitics generated by the logic grammar of this thesis, we shall look at those issues.

I am concerned with the problems of syntax and semantics of clitics in the generation of sentences in the null-subject Romance language Spanish. A problem that has attracted recent research is to find a linguistic theory that demonstrates a basic similarity between the treatment of clitics in the Romance languages Italian, French, and Spanish while still providing a framework that will work for the treatment of anaphora, pronominal reference, and referential expressions in English and the Romance languages. Within GB theory, this comes down to questions of the status of the clitic and its associated NP with respect to the Binding Theory and the Empty Category Principle.



### 1.3.1. Clitics in Spanish

Spanish clitics perform the semantic reference function of pronouns but have the properties of inflections on verbs. The following sentences are examples of those whose S-structure is generated by my logic grammar. Sentence (1.21.a) has an NP (*la niña*) which may be 'pronominalised' with the clitic *la* to give the sentence **Juan la ve**. In the Rio de la Plata dialect (Argentina and Uruguay) the clitic appears even when the NP is overt<sup>18</sup> giving **Juan la ve a la niña**, in which case the clitic chain (*la . . . la niña*) has the property that both elements of the chain are marked (+ovrt). In sentence (1.21.b) the indirect object clitic *le* appears in all dialects, both when the NP is overt and when it is absent (in the latter case providing pronominalisation.) The impersonal *se*-clitic in (1.21.c) is very common in Spanish. No overt subject pronoun, such as *they* in English, appears. Sentences (1.21.d,e) are the active and passive versions respectively of a sentence, both assumed to have the same meaning, and thus the same semantic representation. The benefactive dative construction of (1.21.f) is particularly interesting in that it is idiomatic in English but very common in Spanish. The semantic **experiencer** is rendered in syntax as the dative clitic *me* and the sentence has the impersonal *se*-clitic subject construction.

In current<sup>19</sup> linguistic thought [Jaeggli 82] [Rivas 77], Spanish clitics are Base generated in their appropriate positions. For example, Rivas suggested that all clitic/NP pairs were Base generated in position. A clitic/NP agreement rule then paired clitic and NP. Finally, after all syntactic processes, a rule deletes either the clitic or the NP (or neither).

---

<sup>18</sup>The NP must be definite and animate

<sup>19</sup>Earlier work on clitics had proposed a movement transformation derivation of clitics [Kayne 75] [Quicoli 76]

- (1.21.a) **Juan ve a la niña.**  
*Juan sees the child.*
- (1.21.b) **La criada le escribe a Juanito.**  
*The maid writes to Juanito.*
- (1.21.c) **Se venden diarios.**  
*Newspapers are sold. (They sell newspapers.)*
- (1.21.d) **El médico compró la bencina.**  
*The doctor bought the petrol.*
- (1.21.e) **La bencina fue comprada por el médico.**  
*The petrol was bought by the doctor.*
- (1.21.f) **Se me perdieron las llaves.**  
*My keys got lost on me.*  
*(I lost my keys but it wasn't my fault)*

The NP in a clitic/NP pair such as [la] . . . [al la niña] of (1.23.a) or the NP [a la niña] of (1.21.a) is called the direct object of the verb. The verb is said to govern the direct object. In this case the verb lexically properly governs the NP, which allows the verb to assign Case to the NP position. The notion of proper government is a structural one. For our current purposes, we can use the definition of government in [Chomsky 82a]:

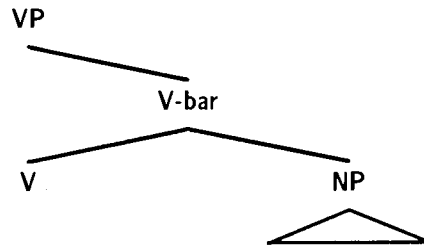
In the structure  $[\gamma \dots \beta \dots \alpha \dots \beta \dots]$

$\alpha$  governs  $\beta$  iff

- (i)  $\alpha$  is an immediate constituent of  $\gamma$
- (ii) where  $\phi$  is a maximal projection,  
if  $\phi$  dominates  $\beta$  then  $\phi$  dominates  $\alpha$ .

So, in the structure (1.22) the verb is an immediate constituent of the V-bar, and the maximal projection VP dominates both the verb V and the NP, and thus the verb governs the NP.

(1.22)



### 1.3.2. Clitics and Empty Categories

Sentences such as (1.23) pose a question about empty categories. In (1.23.a) and (1.23.b) the clitic [la] is doubled with an NP. In (1.23.a) and (1.21.a) it is not. If we assume, as [Rivas 77] did, that an empty category is associated with the clitic (as a result of deletion in Rivas' analysis) in these latter two sentences<sup>20</sup>, then there are consequences to that assumption in GB Theory. I shall examine each of these consequences in turn.

(1.23.a) Juan la<sub>i</sub> ve [a la niña]<sub>i</sub>.

*Juan sees the child.*

(1.23.b) Juan la<sub>i</sub> ve [a ella]<sub>i</sub>.

*Juan sees HER.*

(1.23.c) Juan la<sub>i</sub> ve e<sub>i</sub>.

*Juan sees her.*

The nature of the empty NP category in the non-doubled constructions (1.23.c) has been the subject of some debate. Jaeggli [82] has called this position a **PRO**. This means that the position is, by definition, anaphoric and pronominal, and by the principle C of the Binding Theory (of GB), the position may not be governed. Jaeggli's explanation of this is that the clitic absorbs government, allowing the **PRO** to occupy the position ungoverned. The assumption is that the empty position is an A-position. That is, an NP position that

<sup>20</sup>An empty clitic in (1.21.a).

is Base generated as a subcategorisation position of the verb, and which corresponds to one of the arguments of the semantic proposition. Normally, for an overt NP, this position would absorb Case and  $\Theta$ -role from the verb that governs it. Borer [81] has called this position a properly governed empty category. Again the assumption is that the position is an A-position. Free generation gives a clitic that can be co-indexed at LF with the empty position and filters out those that do not agree. Aoun [85] calls the empty category associated with the clitic a variable and states that the variable is not treated as an argument but as an  $\bar{A}$ -anaphor with no  $\Theta$ -role, and co-indexed with the clitic. The assumption must be here that the clitic position, which Aoun also calls an  $\bar{A}$ -position (see below), must absorb Case and  $\Theta$ -role, and is therefore governed by the verb.

The nature of the clitic position itself is open to debate. This question is examined at some length in Chapter 4. Let me just say here that the clitic has been said to be generated in VP as a separate category [Rivas 77] [Safir 85] and as an adjunct to the verb [Jaeggli 82]. In a refreshing new proposal, [Roberge 85] suggested that French subject clitics were Base generated in AGR under INFL. I suggested, with Sempere in [Brown and Sempere 85], that Spanish object clitics are also Base-generated in INFL, but in a separate inflectional adjunct position called ENCL2. This latter proposal is detailed in Chapter 4. Not much attention has been paid to the question of the nature of the position of the clitic when the clitic does not appear phonetically. The unstated assumption has been that the clitic was never generated and thus no corresponding empty category exists. Nor, presumably, does any structural position for the clitic exist in such cases. The exception has been Rivas, who deleted the clitic. In this case deletion does not involve an empty category. Aoun [85] specifically calls the clitic position itself an  $\bar{A}$ -position in VP, but one that bears a  $\Theta$ -role.

In the manuscript [Hurtado 86] (published posthumously), Hurtado suggested that a Base-generated clitic chain [clitic . . . **pro**] could account for the object clitic doubling and left dislocation sentences in Spanish. I will extend that suggestion in Chapter 2, and Base-generate a clitic chain [clitic . . . **NP**] for all verbal argument positions in Spanish. This is a computational convenience, to avoid the unnecessary generation of all possible combinations of clitics and NPs, most of which would be subsequently ruled ungrammatical. I shall use an OVERT feature<sup>21</sup> for both the clitic and the NP to control which elements of the chain are phonetically empty. This again is a computational convenience, but one which I suggest has linguistic validity.

My suggestion is that clitics chains are Base-generated for all verbal argument positions in Spanish, and that each element of the clitic chain is assigned an overtiness feature [**±ovrt**] in the Base. The emphasis is on each here. The elements of the chain do not necessarily have the same value of the OVERT feature. Thus, a clitic position exists for all cases, whether or not the clitic appears phonetically or not. This is a departure from the unstated assumption of others that no clitic position is generated when the clitic does not appear.

I now introduce my inflectional analysis of Spanish clitics, as a prelude to taking the topic up in detail in Chapter 4.

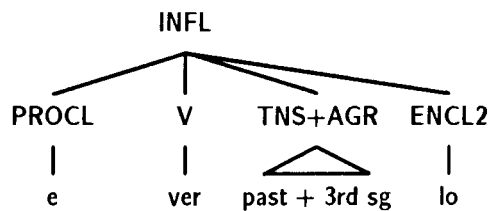
---

<sup>21</sup>As used by [Aoun 85] for A-position anaphors.

### 1.3.3. Base generation of clitics

My theory of the Base generation of clitics in Spanish is introduced in Chapter 4. In it, clitics are viewed as verbal inflections which may move from post-verbal to pre-verbal position. This view contrasts with that of many researchers who have viewed clitics as weak pronouns. The proposed structure of the inflected verb<sup>22</sup> is given in (1.24) with the pre-verbal clitic position (PROCL) and the postverbal clitic position (ENCL2), in addition to the normal subject-agreement position (ENCL1).

(1.24)



My framework of clitic chains uses the OVERT (overtness) feature to provide a computational means of dealing with null-subject, clitic doubling, pronominalisation, and dialectal phenomena. In the future I will use the term clitic chain to refer to a Base-generated set of syntactic constituents, a clitic and an NP, which share the same  $\Theta$ -role, with distinct OVERT features, and which are co-indexed at the level of LF.

(1.25)

**lo**     . . .     **el libro**  
**[+ovrt]**                     **[+ovrt]**

---

<sup>22</sup> I assume that verbs in Spanish are subject to morphological constraints such as those examined for English by [Selkirk 82]. The syntax of words is outside of the scope of this thesis. Thus, the detailed morphological-syntactic structure of the inflected verb does not concern me here.

## 1.4. Summary

This dissertation brings together the concerns addressed very briefly in the preceding sections. That is, the semantic representation of Spanish object clitic constructions and their automatic generation as sentences with a logic grammar of the Constrained Discontinuous grammar type.

In the sphere of Linguistics, the framework is Government and Binding theory in a transformational grammar. A theory of the base generation of clitics is introduced that is fundamentally different from the traditional approach to clitics as a distinct category.

In the computational sphere, it has been found necessary to modify greatly the interpretation of Discontinuous Grammar rules. They are now interpreted as the simultaneous application of two or more rules which share variables for unification and which all must succeed (else the DG rule fails). Active filters, automatic buildup of derivational context, and constraints (after the work of Dahl and Saint Dizier [Dahl and St-Dizier 86]) on that context have been added in to produce a Constrained Discontinuous Grammar

The introduction of the OVERT feature on elements of a clitic chain falls in both the computational and linguistic camps. Without it, automatic generation of the clitic constructions proved impossible. On the other hand, it has independent justification when the semantics of clitic constructions are examined. In order to generate Spanish sentences such as (1.26), where both sentences have the same meaning, a mechanism of feature marking associated with the clitic [la] is required.

- (1.26.a) **Juan la<sub>i</sub> ve [a la criada]<sub>i</sub>.** (Rio de la Plata, other dialects)  
*Juan sees the maid.*
- (1.26.b) **Juan ve a la criada.**  
*Juan sees the maid.*

The area of semantic representation also bridges the computational and linguistic domains. Although an independent semantic representation, as used here, is traditionally outside the sphere of Government and Binding theory, it has been, and is, a part of linguistic inquiry<sup>23</sup>. The examination of the semantic representation of Spanish sentences with clitics has led to a generalisation of the semantics-syntax interface, the Environment of Syntax, in the automatic (computational) generation of sentences.

Finally, the connection between the expressive power shared by Government and Binding theory and Constrained Discontinuous Grammars has been firmly established.

---

<sup>23</sup>For example, the school of Generative Semantics and the school of Montague Semantics.



## Chapter 2

### The Semantics of Spanish Clitics

This chapter deals in detail with the semantic representation of Spanish sentences with object clitics. First the OVERT feature for elements of clitic chains is dealt with in more detail. Then, direct objects, impersonal subjects, and indirect objects will be dealt with in turn.

The semantic representation used in this research is intended to be an interlingual representation for an automatic NL translation system. It is assumed that a (unspecified) NL analysis system is capable of producing the semantic representation which is, in turn, used as an input to the generation grammar. Thus the semantic representation represents a form which a similar generation grammar for another language (English, French, etc.) might use as input. This same semantic representation has been used for a prototype generation grammar [Brown, Pattabhiraman et al 87] [Brown, Dahl et al 86] [Dahl, Brown et al 86] for English. The semantic representation might be used by the response system of a database query system. This could provide, for example, the capability of responses in different natural languages.

The semantic representation used is the Conceptual Graph (CG) of [Sowa 84]. This representation has been found adequate to represent the sentences examined in this research. The CG representation cannot be said to be equivalent to the logical form of [May 84] [May 77] (in linguistics) or [Dahl 77] (in logic grammars) because the scope of NL quantifiers (eg. some, many, all, etc.) is not unambiguous. Nor does it fall under the

heading of word-meaning (although there are provisions in the grammar, not in the semantic representation, for semantic type checking.) Rather, the semantic representation is assumed to be a well-formed representation of a propositional sentence meaning whose conceptual relation names correspond to  $\Theta$ -roles of the arguments of verbs or prepositions, or to the  $\Theta$ -types of nouns in the lexical entries of the lexicon. It is up to the grammar writer and/or the CG composer to ensure this fundamental correspondence. The relationship between the CG representation and the Base grammar is examined in detail in Chapter 3.

## 2.1. Clitic Chains and the OVERT feature

I discussed the notion of clitic chains in Chapter 1. There I showed that there are different views of the nature of clitic constructions. In this Chapter, I explain the view I take of clitic constructions in Spanish. It departs from the current proposals for clitic generation [Jaeggli 82] [Aoun 85] [Hurtado 86] in that I assume that a clitic chain is Base generated for all verbal argument A-positions, whether or not a clitic appears, phonetically. It had previously been assumed that no clitic was generated in these cases. This has been a computational convenience to avoid unnecessary generation of sentences that would otherwise be ruled ungrammatical. By using clitic chains in this way, I have been able to account for the generation of clitic doubling, pronominalisation by clitic, pronominal emphasis, and dialectal variations with one mechanism: the setting of the OVERT feature in pre-lexical processing. I assume that the clitic has no status with respect to the Binding Condition. The treatment of clitics as inflections given in Chapter 4 justifies this assumption.

A syntactic chain was defined<sup>24</sup> to be a set of syntactic constituents which have the

---

<sup>24</sup>See for example, [Safir 85].

same  $\Theta$ -role and which are therefore co-indexed at the level of LF. A clitic chain is, then, a chain including a clitic and an A-position/ $\bar{A}$ -position<sup>25</sup>:

$$(2.1) \quad \text{clitic}_i \quad . . . \quad \text{A-position}_i$$

**Hypothesis: A clitic chain is Base-generated in Spanish for each and every A-position.**

I give each element of the clitic chain a distinct OVERT feature. This is a departure from accepted syntax. It is done as a computational convenience because the setting of these two features in the pre-lexical processing allows the grammar to handle the pronominalisation, clitic doubling, emphasis, and dialectal variations with one mechanism (and avoids unnecessary backtracking in the logic grammar). The OVERT feature expresses the phonological overtness of an element of the clitic chain, and is set to [+ovrt] or [-ovrt] by semantic conventions (pronominalisation, emphasis, and dialect) associated with specific constructions. The setting of OVERT to [+ovrt] means that the element is expressed phonologically. Similarly, [-ovrt] means that the element is not expressed phonologically.

$$(2.2) \quad \begin{array}{ccc} \text{clitic}_i & . . . & \text{A-position}_i \\ [\text{OVERT}] & & [\text{OVERT}] \end{array}$$

**Hypothesis: Both the clitic and the A-position of a clitic chain have a distinct OVERT feature.**

---

<sup>25</sup>It is assumed for the moment that the position in question is an A-position (Argument position in syntax, thus a direct-object position, an indirect-object position, or a subject position). This assumption is brought into question in the section on possessive datives where the possibility is raised that an  $\bar{A}$ -position or a determiner with a  $\Theta$ -role may be involved.

I have found that, using this hypothesis for the generation of the Spanish constructions investigated, at least one element of a clitic chain with an A-position must have the OVERT feature set to [+ovrt]. In some cases, as will be seen, both elements may be [+ovrt]. We now look at the mechanics of the use of this feature for the clitic constructions of Spanish mentioned above.

**Hypothesis:** At least one element of a clitic chain in a tensed construction must have the OVERT feature set to [+ovrt]<sup>26</sup>

## 2.2. Direct Object Clitics

The simple sentence (2.3) with the clitic chain [e] . . . [el libro] is a case of the object clitic chain having one element (the A-position NP, [el libro]) marked [+ovrt] and the other element (the clitic position [e] with no phonological realisation) marked [-ovrt].

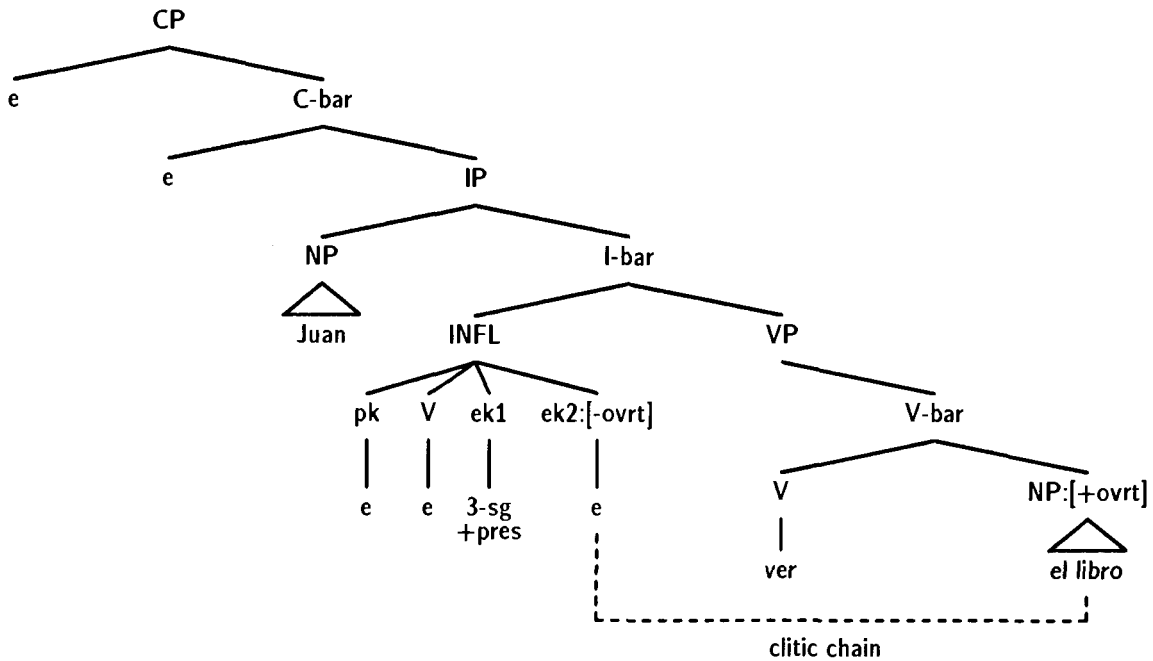
(2.3) Juan e<sub>i</sub> lee [el libro]<sub>i</sub>.  
*Juan reads the book.*

This may be taken as the default setting for direct object clitic chains. The clitic chain in the D-structure is illustrated in (2.4), where the proclitic, enclitic1, and enclitic2 positions are marked as pk, ek1, and ek2 respectively.

---

<sup>26</sup>For infinitival constructions, where the subject A-position is a *pro* (empty pronominal) the situation is not clear. As stated in Chapter 1, it is not the purpose of this thesis to deal with infinitival constructions. The question is of interest if the extensibility of my clitic chain paradigm is to be considered. One possibility is that a clitic chain with both elements [-ovrt] is Base generated for the subject position. Chain composition (as proposed in *Barriers* [Chomsky 86]) could then occur between the clitic chain with both elements [-ovrt] and another clitic chain in the matrix clause which has one element set to [+ovrt]. That is, exactly the same type of mechanism of chain composition applies as that hypothesised in *Barriers* for the empty operator chains of parasitic gap constructions. This remains to be investigated.

(2.4) **Juan lee el libro** (D-structure)  
*Juan reads the books.*



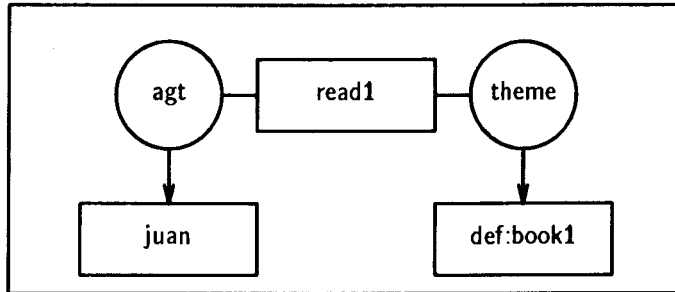
### 2.2.1. Conceptual graph representation

The conceptual graph (CG) representation of the meaning of (2.3) given in (2.5) illustrates the fundamental structure of the CG representations used. The predication<sup>27</sup> **read1** corresponds in this case to a lexical entry for the Spanish verb **leer**. The predication has two relations associated with it: the **agt** and **theme** relations which correspond to the external and internal arguments respectively of the lexical entry. Those

<sup>27</sup>For ease of understanding, English is used as a meta-language of the semantic representations. Thus **read1** is taken to incorporate the normal sense of the verb **read** in English, whereas **read2** would be a different sense of the verb such as the British sense of *to study*.

relations relate the predication **read1** to the concepts labelled **juan**<sup>28</sup> and **def:book1**. The concept **def:book1** indicates that this is a definite reference to the concept **book1** (and not a generic reference, which would be labelled **indef:book1**).

- (2.5) **Juan lee el libro.** (Semantic representation.)  
*Juan reads the book.*



Whether an A-position in the syntax is to be pronominalised or not will not be indicated in the CG. It is up to the procedures consulting the EOS to infer pronominalisation. The CG will always have a label for a concept, whether or not the concept is a predication concept. Pronominalisation will take place, for example, if the concept has previously been mentioned in the discourse, and if there is no possible ambiguity in the pronominalisation. Thus, the consultation with the EOS for the NP involves (among other things) setting the OVERT features of the clitic and the NP. Chapter 3 deals with this question in more detail.

---

<sup>28</sup>The proper name **juan** should perhaps be expressed as **def:juan** in the semantic representation. This would reflect the fact that a reference to a proper name is a definite, and not an indefinite, reference. This treatment would require a flagging of those concepts which are proper names in order to avoid the generation of a definite article in most cases. Exceptions exist, as in the sentence: *Have you seen the Jones today?*. On the other hand, reference may be indefinite, as in: *Is there a Peter in the class?*. I do not address these problems in this dissertation.

### 2.2.2. Emphatic and nonemphatic pronominal NPs

When a direct object is pronominalised in Spanish, the NP position<sup>29</sup> becomes [-ovrt] and the clitic position in the clitic chain must be [+ovrt]. So, when the direct object NP **Juanita** in (2.6.a) is pronominalised we get (2.6.b). Direct object clitics may optionally co-occur with their co-referential direct object NPs (clitic doubling). When the direct object NP is pronominalised overtly in such a doubling construction, the co-referential direct object clitic remains overt as in (2.6.c).

- (2.6.a) **La funcionaria [e]<sub>i</sub> llamaba [a Juanita]<sub>i</sub>.**  
*The clerk called Juanita.*
- (2.6.b) **La funcionaria [la]<sub>i</sub> llamaba e<sub>i</sub>.**  
*The clerk called her.*
- (2.6.c) **La funcionaria [la]<sub>i</sub> llamaba [a ella]<sub>i</sub>.**  
*The clerk called HER.*

However, in sentence (2.6.c) the *doubled* pronominal-NP [**a ella**] is emphatic. Non-emphatic pronominal-NPs are not overt, but are expressed through the clitic. Sentence (2.6.b) is the non-emphatic form of (2.6.c).

An operator relation (**op**) for **emphasis** in the semantic representation (2.10) is realised in the normal dialect of Spanish with a [+ovrt] feature in the A-position when the A-position (**a ella**) of the clitic chain in (2.7.a), is also marked syntactically as [+pronominal,+emphasis]. When the pronominal A-position is not marked for emphasis, the corresponding chain features are as in (2.7.b) with the A-position marked [-ovrt].

---

<sup>29</sup>Note that in (2.6.a) the phrase [**a Juanita**] is an NP, not a PP. In Spanish, a *personal* [a] is added to animate object NPs. There is some evidence both in Spanish and in French to suggest that this is a *late* phenomenon.

(2.7.a)

<b>clitic</b>	. . .	<b>A-position</b>
[+OVRT]		[+OVRT.+PRONOMINAL.+EMPHASIS]

(2.7.b)

<b>clitic</b>	. . .	<b>A-position</b>
[+OVRT]		[-OVRT.+PRONOMINAL.-EMPHASIS]

A similar situation pertains for topicalised pronominal direct object NPs<sup>30</sup> as in (2.8).

(2.8.a)        **[A Juanita]<sub>i</sub>; la funcionaria [la]<sub>i</sub>; llamaba.**  
*JUANITA, the clerk called her.*

(2.8.b)        **[A ella]<sub>i</sub>; la funcionaria [la]<sub>i</sub>; llamaba.**  
*HER, the clerk called.*

I suggest that the operator relation for **emphasis**, at the level of semantic representation also exists for (2.8) and that the respective chain marking is as found in (2.9). The sole difference from (2.6.c) being that the A-position may be marked  $\pm$ -pronominal. A conclusion here is that the features of the clitic chain of (2.9.b) (shown here at D-structure, before the topicalisation move- $\alpha$ ) must force topicalisation orthographically, although not necessarily phonologically.

(2.9.a)

<b>clitic</b>	. . .	<b>A-position</b>
[+OVRT]		[+OVRT.-PRONOMINAL.+EMPHASIS]

(2.9.b)

<b>clitic</b>	. . .	<b>A-position</b>
[+OVRT]		[-OVRT.+PRONOMINAL.+EMPHASIS]

I conclude that the Conceptual Graph representation of such sentences includes an **emphasis** relation on the concept referred to by the R-expression in A-position. This is

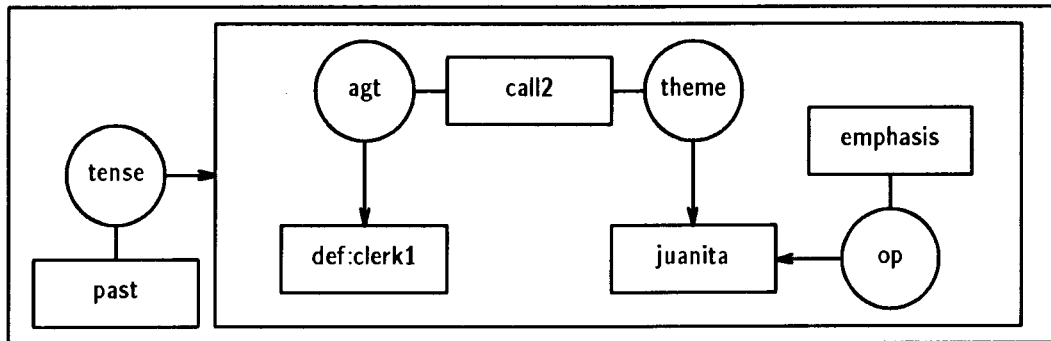
---

<sup>30</sup>I do not deal with left dislocated NPs in this dissertation. They are NPs which are Base generated in topicalised position but are followed by a resumptive pronoun (unless that is impossible). Thus: **[A Juanita]<sub>i</sub>; la funcionaria [la]<sub>i</sub>; llamaba [a ella]<sub>i</sub>;**



necessary, over and above any inference of pronominalisation because, as the examples above demonstrate, a pronominalisation feature in itself is not enough to account for these emphatic and non-emphatic constructions.

(2.10) [A Juanita]<sub>i</sub>; la funcionaria [la]<sub>i</sub>; llamaba.



One note here The operator relation on the concept **emphasis** is not one of the argument-relations of the predication **call2**. We will find similar situations when dealing with dative clitics

### 2.2.3. Dialectal direct object doubling

Some dialects, notably the Rio de la Plata dialect of Argentina and Uruguay, require a doubling clitic for specific animate direct objects.

(2.11.a) El general  $los_i$  mira a [los soldados]<sub>i</sub>.

*The general looks at the soldiers.*

(2.11.b) La gente  $las_j$  trae [las niñas]<sub>j</sub> a la plaza.

*The people carry the little girls into the plaza.*

The clitic chain for the direct object in this dialectal case is therefore marked as (2.12):

(2.12)

clitic	. . .	A-position
[+OVRT]		[+OVRT.ANIMATE]

There is no emphatic connotation in this construction. Emphasis is obtained through phonetic emphasis or through topicalisation. Thus, (2.13) represents the clitic chain for (2.11.b)

(2.13)

<b>las</b>	. . .	<b>las niñas</b>
[+OVRT]		[+OVRT.-PRONOMINAL.-EMPHASIS]

This construction provides evidence that the features [ $\pm$ pronominal, $\pm$ emphasis] are not in themselves enough to explain the synchronic dialectal data. The OVERT feature has been introduced to solve this problem in a convenient way for the automatic generation.

#### 2.2.4. Topicalised direct objects

Topicalised direct object NPs require an overt co-referential direct object clitic as we saw above in (2.8). It is possible that binding of the clitic to the NP takes place before extraction of the NP to a topicalised position (through a *move- $\alpha$*  transformation). Such an approach has led to divergent views on the following syntactic questions<sup>31</sup>.

- 1 What is the nature of the empty category associated with the clitic in the non-doubled constructions?
- 2 What is the argumental status of the doubled NP in doubled constructions?
- 3 What is the nature of the empty category associated with the clitic and the extracted NP<sup>32</sup>?

For example, [Jaeggli 80] considers the empty category associated with the clitic in the non-doubled construction to be an ungoverned PRO<sup>33</sup> while [Borer 81] considers it to be a

---

<sup>31</sup>These are the general form of the questions found in [Hurtado 86].

<sup>32</sup>Assuming that the extracted NP leaves behind some empty category.

<sup>33</sup>The empty anaphoric pronominal operator. That is, a variable with pronominal marking, and no Case.

properly governed empty position (**e**) and [Aoun 81] considers it to be a non-argumental-anaphor ( $\bar{A}$ -anaphor) with no  $\theta$ -role. On the other hand, the empty category associated with a clitic and an extracted NP is considered by Borer to be a trace (**t**) and by Aoun to be an  $\bar{A}$ -anaphor with a  $\theta$ -role. All three researchers consider a non-extracted, doubled NP to be in argumental position.

My position is that the empty category associated with the clitic in the non-doubled non-topicalised constructions is a lexically properly governed<sup>34</sup> A-position. It is not empty as a result of a movement transformation, but simply as a result of being marked *-ovrt*. The empty A-position could therefore be taken to be a **pro** (not **PRO** as Jaeggli suggested) but properly governed (in opposition to Jaeggli but in agreement with Borer) chained to an inflection clitic position in exactly the same way as the empty pronominal subject A-position is taken to be **pro** chained<sup>35</sup> to the AGR+TENSE inflection of verbs. This working hypothesis will be expanded upon in Chapter 4. As for the doubled constructions, the doubled NP is in A-position, again lexically properly governed. In the case of topicalised NPs, the trace (in A-position) of the topicalised NP (in  $\bar{A}$ -position) surely remains lexically governed. This analysis differs from the above-mentioned analyses in that it differentiates between the topicalised and non-topicalised doubled constructions.

---

<sup>34</sup>As defined in Barriers [Chomsky 86].

<sup>35</sup>Called *identification* by Jaeggli and Hurtado.

## 2.3. Impersonal-se Constructions

The impersonal construction (2.14.a,b,c) in Spanish should be carefully distinguished<sup>36</sup> from the intransitive (2.14.d) construction (see [Westfal 79] [Knowles 74]).

- (2.14.a) **Se venden diarios allá.**  
*They sell newspapers over there.*  
*(Somebody unspecified sells newspapers over here).*  
*(Newspapers are sold - by somebody unspecified - over there).*
- (2.14.b) **Se abrieron las puertas.** (Standard Spanish)  
*The doors were opened (by somebody unspecified).*
- (2.14.c) **Se abrió las puertas.** (Some dialects of Spain)  
*The doors were opened (by somebody unspecified).*
- (2.14.d) **Las puertas se abrieron.**  
*The doors opened.*

I assume the impersonal-se construction has an **agent**  $\Theta$ -role<sup>37</sup> which is an unspecified concept in the semantic representation. I call this concept **unspec**. The unspecified agent is realised in the Base generation as a clitic chain. The subject A-position is assumed<sup>38</sup> to be **pro**, that is, a phonetically unrealised pronominal element chained to the subject clitic **se**. As a pronominal, **pro** is free in its governing category (according to principle B of the Binding Condition of GB theory).

- (2.15)
- |            |       |           |
|------------|-------|-----------|
| <b>pro</b> | . . . | <b>se</b> |
| [-OVRT]    |       | [+OVRT]   |

---

<sup>36</sup>I have decided to follow the analysis of [Westfal 79] in the analysis of these sentences.

<sup>37</sup>Another possibility is to follow [Chomsky 82a] and call this a **subject-theta-role**. This would present difficulties in the pre-lexical processing of the semantic representation in cases where the lexical entry for a verb specified an agent and the semantic representation specified a **subject-theta-role**. I have decided to use **agent** on the purely pragmatic ground of computational simplicity.

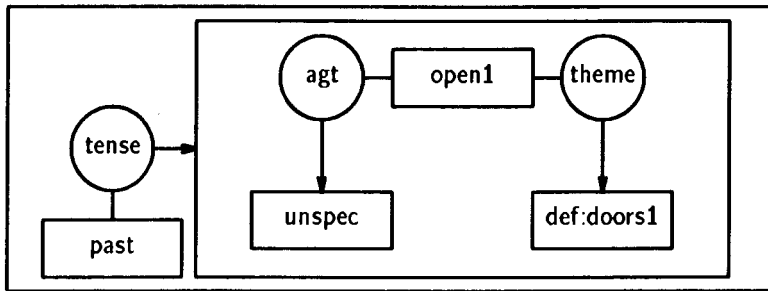
<sup>38</sup>This follows from the discussion of empty categories in *Some Concepts and Consequences of Lectures on Government and Binding* [Chomsky 82b].

It should be noted that verb agreement<sup>39</sup> is a 'late' phenomenon in these constructions.

Thus the semantic representation of both (2.14.b) and (2.14.c) is (2.16):

(2.16.a) **Se abrieron las puertas.**

(2.16.b) **Se abrió las puertas.** (Some dialects of Spain)



The marked construction is (2.14.b) where object-verb agreement occurs. The unmarked (2.14.c) has the default third person singular verb agreement.

The intransitive construction has markedly different semantics. Note that the predication in the semantic representation (2.17) is **open2**, a different predication than that in (2.16). Corresponding lexical entries will have different argument structures.

## 2.4. Indirect Object Clitics

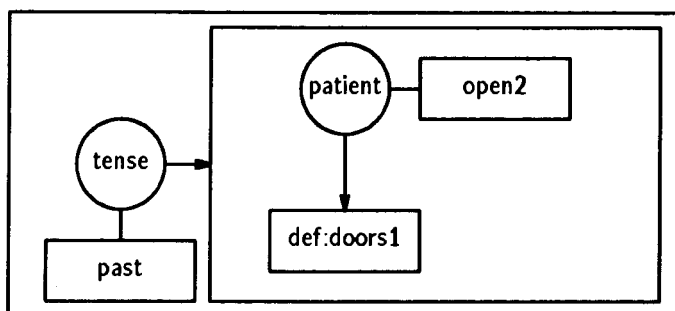
Spanish is rich in dative constructions which are often used to express personal involvement in a situation. In addition to the "normal" dative sense of *to* or *for*, Spanish uses the dative for focusing attention on an individual or the cause of an event.

Studies of Spanish grammar use terms such as *ethical dative*, *benefactive dative*,

---

<sup>39</sup>Backwards agreement may apply here. That is, for transitive verbs, agreement may hold between the verb and the complement of the verb.

## (2.17) Las puertas se abrieron.



*datives of interest*, and others. The use of these terms is not standardised. Some authors have used the terms to distinguish  $\Theta$ -roles of datives, others use them to distinguish syntactic behaviour. In this section we are discussing the semantics of clitics, thus the use of these terms here is to distinguish between the different semantic uses of dative clitic constructions.

As a starting point I accept the judgement of the Real Academia Española [Bello y Cuervo 47] that dative of interest, possessive dative, and ethical dative have distinct semantics. To this list, I add the benefactive dative and show in this section how I feel the semantics of this construction differ from the others.

#### 2.4.1. Dative clitic doubling

In normal usage the dative object of a verb is doubled with a clitic. That is the clitic chain is marked with [+OVRT] at both the clitic and the A-position.

(2.18)

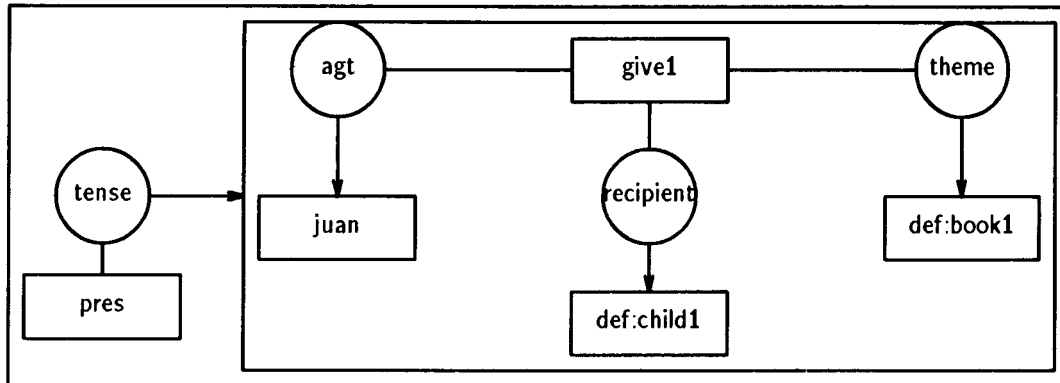
<b>clitic</b>	. . .	<b>A-position</b>
[+OVRT]		[+OVRT]

Sentences such as (2.19), where the clitic *le* is doubled with the indirect object *al niño*, are normal usage in all dialects.

- (2.19) Juan le<sub>i</sub> da el libro [al niño]<sub>i</sub>.  
*Juan gives the book to the child.*

The thematic role of the indirect object may be RECIPIENT.

- (2.20) Juan le da el libro al niño.



#### 2.4.2. Possessive dative

The *possessive dative* is a common usage which implies possession of something, often the direct object, by a dative clitic. In these constructions the dative clitic is required and is usually not doubled with a dative NP (although it is possible as in (2.21.c)).

- (2.21.a) Le cerraron los ojos.  
*They closed his eyes.*
- (2.21.b) No le pagaron el sueldo.  
*They did not pay him his wages.*
- (2.21.c) No le<sub>i</sub> pagaron el sueldo [a Juan]<sub>i</sub>.  
*They did not pay Juan his wages.*

Parts of the body are not usually referred to by a possessive in Spanish<sup>40</sup>. Sentence (2.21.a) is typical of this construction in that a definite article (*los* = *the*) is used instead

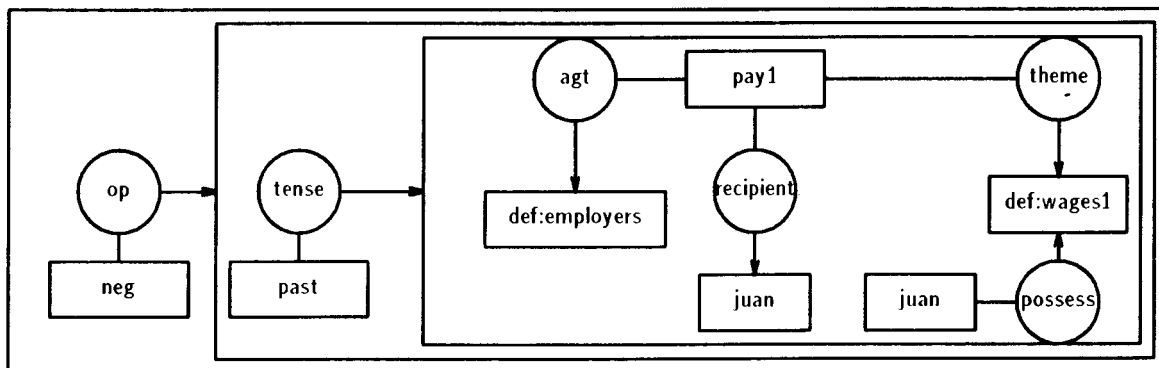
<sup>40</sup>This is typical of other languages as well. North Germanic languages normally use this construction.

of a possessive adjective (*sus* = *his*) and a dative clitic (*le* = *to him*) is used to indicate whose eyes. In (2.21.b) a definite article, *el*, is used as a specifier for *wages* and a dative clitic, *le*, indicates whose wages.

There appears to be no differentiation made with regard to the argumental status or presence of the dative object.

In (2.22)<sup>41</sup> the dative clitic is chained to the recipient A-position which is marked [-ovrt]. On the other hand, the dative clitic is also co-referential<sup>42</sup> with the determiner [*el*] (of the direct object NP [*el sueldo*]) which carries the possess  $\Theta$ -role. (Note: (2.22) assumes that the recipient is pronominalised.)

(2.22) **Los amos no le pagaron el sueldo.**  
*The bosses did not pay him his wages.*



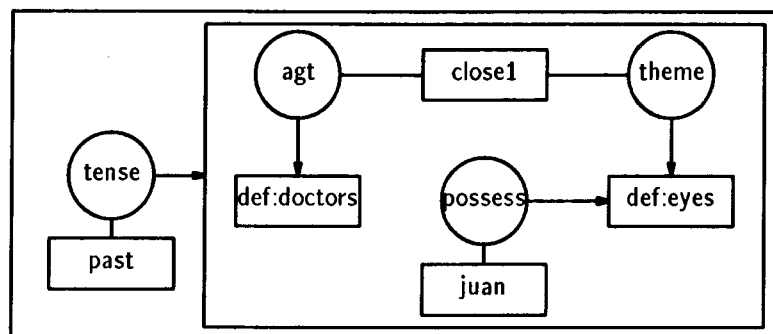
<sup>41</sup>In (2.22) I have assumed that the negation has wider scope than the tense. It is possible that the scope of these two operators should be reversed. Alice ter Muelen of the University of Washington has suggested that an episode is dominated by tense in *The Semantic Representation of Tense and Aspect*, a lecture delivered at Simon Fraser University, 12 May, 1987. This same point could be made for (2.25-6), (2.28-9), and (2.31-2).

<sup>42</sup>If we assume that the possessive PP [*de Juan*] is suppressed in the NP [*el sueldo de Juan*] (the pay of Juan) we can assume a co-referential possessive dependency between the article *el* and the empty PP. The dative clitic is thus actually co-referential to the whole *el* . . . PP dependency.



However, in (2.23) the dative clitic *le* can only be co-referential to the determiner [*los*] (of the NP [*los ojos*]), an  $\bar{A}$ -position, because the lexicon entry for the predication *close1* will have only the  $\Theta$ -roles **agent** and **theme**, and no **recipient** argument.

- (2.23) **Los médicos le cerraron los ojos.**  
*The doctors closed his eyes.*



As a working hypothesis, I assume that the possessive dative clitic is co-referential to the determiner of the direct object NP. This raises a question of whether or not the **recipient** relation actually exists in the semantic representation of (2.22). If so, as I assume, the dative clitic appears to have two semantic<sup>43</sup> functions: **recipient** (a  $\Theta$ -role) and **possessor**.

The Theta Criterion does not allow an argument to have two  $\Theta$ -roles<sup>44</sup>. The definition of syntactic chains stated that the elements of a chain have the same  $\Theta$ -role [Safir 85]. This leads me to the conclusion that there are two levels of co-reference involving the

<sup>43</sup>Stowell [1981] discussed arguments with two  $\Theta$ -roles. In his dissertation, the pair of  $\Theta$ -roles assigned to an A-position becomes inseparable in the syntactic component.

<sup>44</sup>The Theta Criterion of Government and Binding Theory states that each argument bears one and only one  $\Theta$ -role

dative clitic *le* in (2.22) and that these are not *composed*. The two levels have distinct  $\Theta$ -roles. One is the co-referential **possess** dependency between the dative clitic and the definite article. The other is the clitic chain between the dative clitic and the **recipient** NP. In (2.23) there is only one level of co-reference, reflecting the absence of the **recipient** argument.

The possession relation in such semantic representations must be interpreted during Base generation as requiring only a definite article specifier for the possessed object. The possessor must be realised as at least a dative clitic (if not a full PP), but not as a possessive adjective. In addition, there is the possibility that a clitic chain also exists involving the dative clitic with an A-position dative (recipient) NP.

### 2.4.3. Dative of interest

The *dative of interest* is a construction in which the dative clitic serves to involve the referent of the position chained to the clitic in some special way. The referent of the clitic *me* in (2.24) is not normally an argument of the verb *comer* (to eat). I postulate an **interest** relation between the proposition of the sentence and the concept of interest (the speaker in (2.24)).

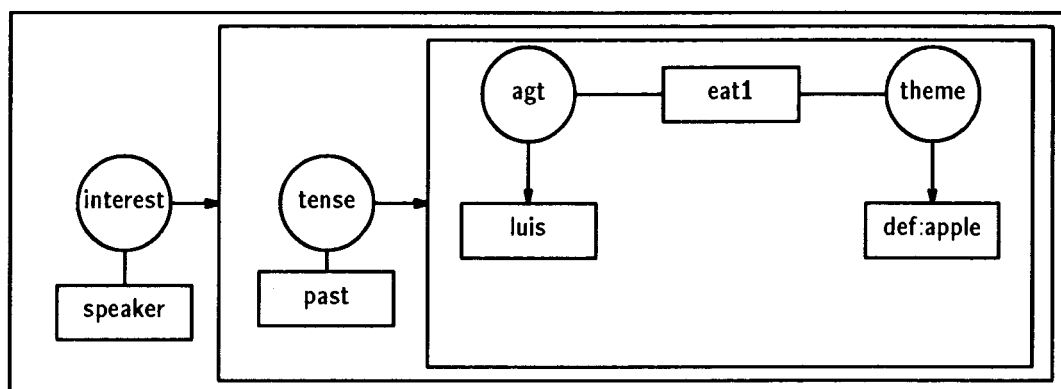
(2.24) *Luis me comió la manzana.*      *from: Strozer 3.62.b*  
*Luis ate the/my apple up on me.*

Thus in the CG (2.25) the **interest** relation is between the proposition, *Luis ate the apple*, and the concept of interest, the *speaker*.

Syntactically in Spanish, this **interest** relation is realised with a

dative clitic. In most cases, this dative clitic is chained to a [-ovrt] NP. Because this

(2.25) Luis me comió la manzana.



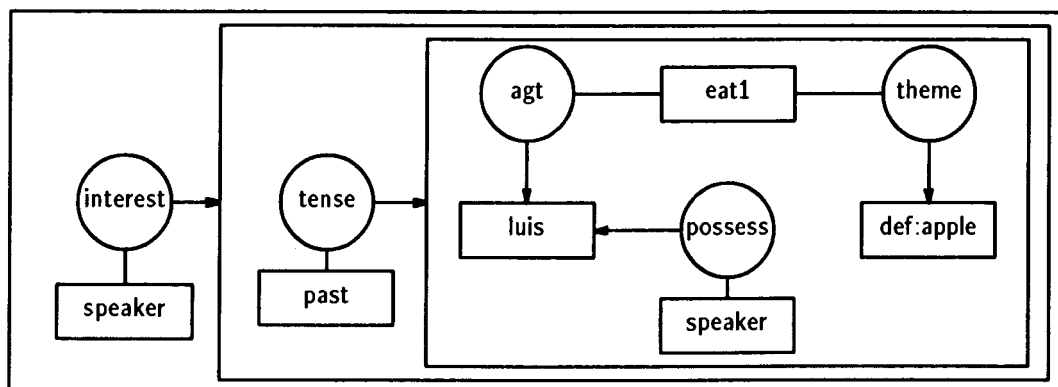
NP is not part of the normal argument structure of the verb. I speculate that the NP is in  $\bar{A}$ -position

In (2.25), the same sentence without the clitic [me] means *Luis ate the apple*, whereas with the clitic an emphatic or dramatic character is added. In the English translation, the PP [up on me] is added to the basic structure of the sentence. Thus (2.24) translates as *Luis ate the apple up on me*. An argument has been made [Strozer 76] that the clitic *me* is a reflexive pronoun. It is difficult to accept that it is merely reflexive. One might attribute both dative and reflexive properties to the clitic<sup>45</sup>.

It is possible to interpret the definite article *la* as a possessive article in uses of the dative of interest as illustrated in (2.26). The possessor is assumed to be the person involved in the *interest* relationship.

<sup>45</sup>Sentence (2.24) may have another meaning in certain contexts: *Luis ate the/my apple up for me (because I asked him to)*. This is difficult to handle. A possibility is to use an [+OBLIGATION] relation at the level of semantic representation to capture this meaning. The sense is that it is *in my service* or *in my interest*.

- (2.26) Luis me comió la manzana.  
*Luis ate my/(?\*the) apple up on me.*



In summary then, the dative of interest appears to have an **interest** (and possibly an **obligation**) relation at the level of semantic representation. It is important to note that the **interest** relation is outside of the set of internal and external arguments of the lexicon entry for the predication **eat1**. Thus, in the example (2.24) the external argument has the  $\Theta$  role **agent** and the internal argument has the  $\Theta$ -role **theme**. These  $\Theta$ -roles are of course represented in the CG by the relations **agt** and **theme**, and the relation **interest** is additional to those demanded by the lexicon. **Hypothesis: An extra internal argument is added by prelexical processing for datives of interest. The argument position is marked [-ovrt] and is chained to the dative clitic marked [+ovrt].**

It is possible that the [-ovrt] NP chained to the dative is in  $\bar{A}$ -position. If that should prove to be the case, then no prelexical processing is necessary and the NP would presumably be adjoined to VP or IP in D-structure. At this time I know of no tests to determine the argumental status of this [-ovrt] position.

As we have seen earlier, it is possible for the dative clitic to also be co-referential to a

"possessive" definite article. Again, It is assumed that the two levels of co-indexation are not composed as they have different  $\Theta$ -roles.

#### 2.4.4. Ethical dative

To denote a personal effect on the speaker Spanish uses the ethical-dative. The meaning is one of moral affront to the speaker. Thus, in (2.27.a) *Don't get carried away on me*, the speaker is saying that it is a personal affront that you are getting heated (in debate). Similarly in (2.27.b) the speaker is saying it is an emotional affront that her/his best friend died.

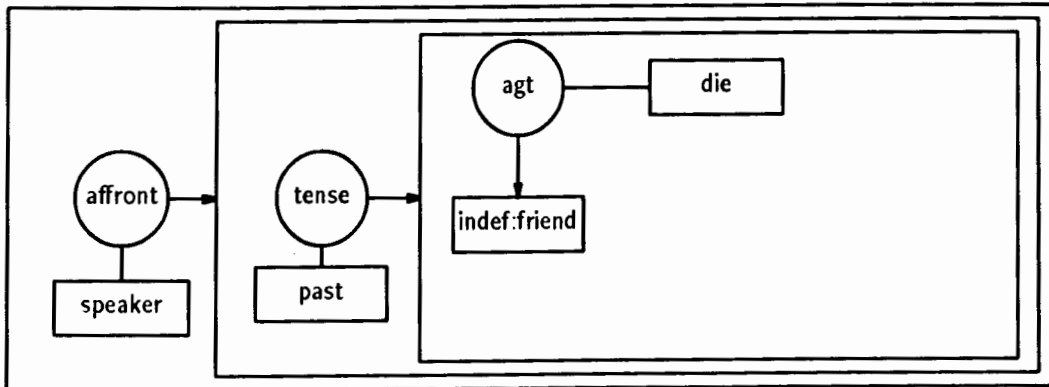
- (2.27.a) No se me acalore.  
*Don't get angry on me.*
- (2.27.b) Se me murió mi mejor amigo.  
*My best friend died on me.*
- (2.27.c) No te me vayas.  
*Don't go away on me.*

If the affront is treated as a relation in the semantic representation outside of the normal argument structure associated with the lexical entry for the predication, the structure is found to be much the same as that of the dative of interest constructions. The dative clitic is chained to an [-ovrt] NP in the Base generation. Again, the argumental status of this NP is not clear. I assume, as in the dative of interest construction, that this is an extra argument added by prelexical processing.

Thus in (2.28) the lexicon specifies only an **agent**  $\Theta$ -role for the predication **die**. The **affront** relation in the semantic representation is extra to the expected set of arguments. The realisation of this extra argument is syntactically a dative [-ovrt] NP position chained with the clitic **me**. Again, the argument status of this [-ovrt] NP is open to debate.

**Hypothesis: An extra internal argument is added by prelexical processing for**

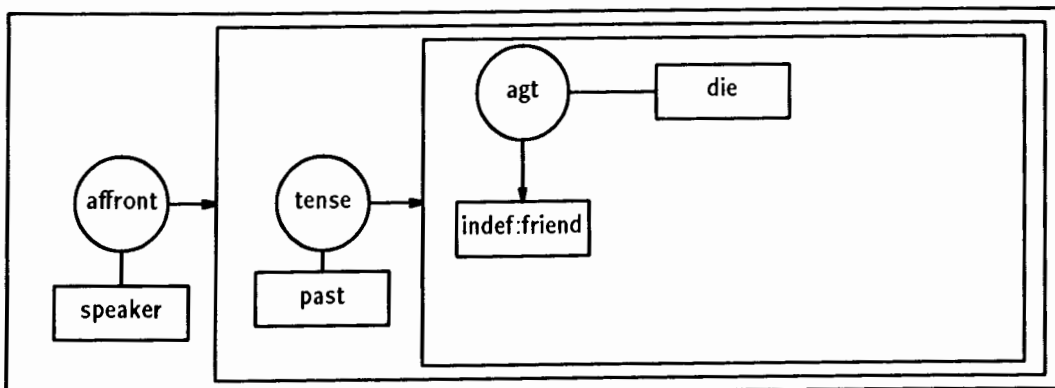
- (2.28) **Se me murió un amigo.**  
*A friend died on me.*



ethical datives. The argument position is marked [-ovrt] and is chained to the dative clitic.

As with the dative-of-interest constructions, the **possess** relation in the semantic representation is possible and may or may not be expressed in the syntax as a possessive adjective. In this case also, the possessor is the person involved in the affront as is seen in (2.29).

- (2.29) **Se me murió un/mi amigo.**  
 (?\*A)/*My friend died on me.*



It is again assumed that the dative clitic may also be chained to the "possessive" article and that the two chains are not composed because of the different  $\Theta$ -roles.

#### 2.4.5. Benefactive dative

I will use the term *benefactive dative* to describe those datives which occur in impersonal constructions to imply that whatever happened, happened accidentally or was not the fault of the speaker. This may also be thought of as a dative of unexpected occurrence. This always involves an impersonal-*se* construction of the verb.

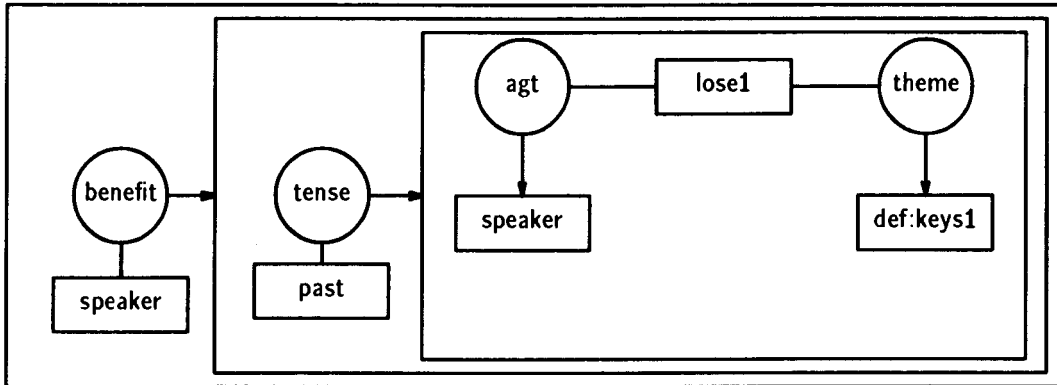
- (2.30.a) **Se me quedaron las cosas en casa.**  
*The/my things were left behind at home.*
- (2.30.b) **Se me perdieron las llaves.**  
*The/my keys got lost on me.*
- (2.30.c) **Se nos pasó el tiempo.**  
*Time passed on us.*
- (2.30.d) **Se me rompió el vaso.**  
*The vase broke on me.*

The sense of (2.30.a) *The/my things were left behind at home*, or more literally *My things stayed at home on me*, is that it was unexpected or accidental that my things remained at home. Similarly with (2.30.b), *The/My keys were lost on me*, the meaning is that it was not my fault or was unexpected.

The sense however is a particular one that expresses some benefactive intent to the animate topic of the sentence. For this reason I take the benefactive dative to be a distinct semantic phenomenon. As discussed above there are slightly different senses of the benefactive intent in the sentences, however I feel there are sufficient grounds to tentatively postulate a **benefit** relation at the level of semantic representation as in (2.31).

In these cases some prelexical processing has taken place. The **agent** has become,

- (2.31) **Se me perdieron las llaves.**  
*The keys got lost on me.*



syntactically a dative internal argument, and an extra unspecified external argument has been added that is syntactically realised as an impersonal-se construction.

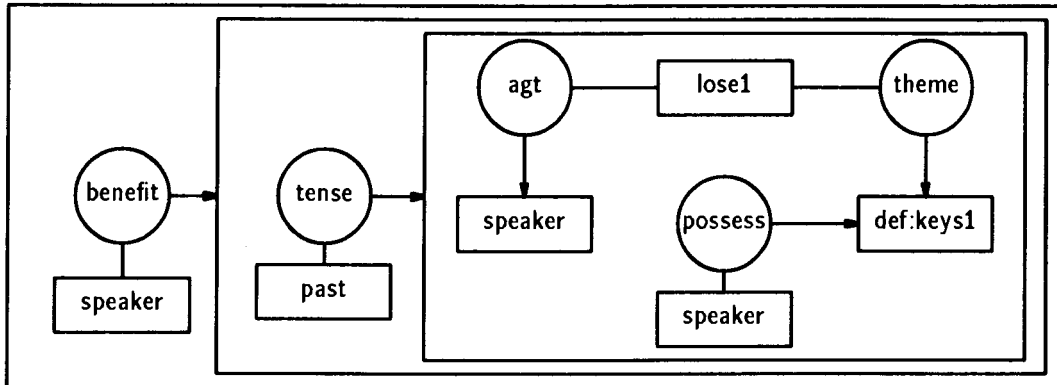
**Hypothesis:** In the benefactive dative construction, prelexical processing introduces a (new) 'experiencer' external argument and the external 'agent' argument becomes an internal argument.

The dative clitic in these expressions may also be interpreted as a possessive dative. In this case also, the possessor is the same person as the person involved in the benefit relation.

Once more then, the dative clitic may be an element of two chains. One involves the **possess**  $\Theta$ -role, and one involves a [-ovrt] NP position with the **agent**  $\Theta$ -role in this case.



- (2.32) **Se me perdieron las llaves.**  
*My keys got lost on me.*



## 2.5. Summary

This Chapter has examined in detail the semantic representation of object clitic constructions in Spanish. Semantic representations have been suggested. These constructions and corresponding syntactic structures have been discussed. I have assumed an analysis of Base generation of all A-positions as involving clitic chains. I have also assumed that both the clitic and the A-position have a distinct OVERT feature. This analysis of Base generation has been adopted for computational reasons: The desire to have one computational mechanism to handle pronominalisation, emphasis, and clitic doubling. It may be that the analysis will prove to be useful in the linguistic arena, as it does not violate Theta Theory, Government, or Binding Theory in generation of D-structure. This depends on whether the hypothesis of clitics as inflections (presented in Chapter 4) and the assumption that clitics have no status with respect to the Binding Theory is accepted.

Semantic representations have been proposed for the syntactic structures of the Spanish constructions examined.

A set of pre-lexical processes that amount to building new argument structures for verbs has been proposed for the set of dative clitic constructions. The implementation of such pre-lexical processes in the logic grammar for Spanish generation is discussed in Chapters 3 and 5

## Chapter 3

### Syntax, Semantics, and the Environment of Syntax

#### 3.1. Introduction

The theory of automatic NL generation presented here has its roots in the transformational-generative theory of language commonly called Government and Binding (GB) theory. The main point is that an interface is provided between the grammar phrase structure rules and the environment of syntax (EOS) (of which the semantic representation is only one part) and that this interface takes place at the maximal projection<sup>46</sup> (XP) level. The EOS represents the environment of generation local to one sentence. As presented here it contains the following components.

- An augmented lexicon (a local lexical context).
- A categorial type (a local syntactic context).
- A semantic representation for one sentence (a local semantic context).
- A semantic-type hierarchy (a local pragmatic context).
- An R-expression<sup>47</sup> 'stack' (a local discourse and sentential context).

A set of procedures for consulting the EOS exists and is used by the syntactic component of the generation grammar. These procedures must be specified in terms of the grammatical features required by the syntactic component. One such set of procedures will be discussed in detail in Section 3.4.

---

<sup>46</sup>A maximal projection is a major phrasal category, that is one of { NP, VP, PP, AP, CP, IP }.

<sup>47</sup>An R expression is a referential expression: an NP referring to some concept.

The EOS may not be complete, but is I believe a step in the direction of a universal model (and one that has been found adequate and useful for the constrained problem of NL generation examined in this dissertation). Its constituents have been found necessary, but they probably do not represent a complete local environment (or *context*) for syntax. The theory presented here provides a principled account of the interface between the environment of syntax (EOS) and the syntax of the generative grammar.

It is assumed that a sentence is generated from a semantic representation which is complete and correctly encodes the meaning of the target sentence(s). The semantic representation chosen for the implementation of this work is the *conceptual graph* introduced by Sowa, discussed at length in [Sowa 84], and used and extended in [Brown, Pattabhiraman et al 87] [Foo and Sowa 87]. Its usefulness for analysis is currently being explored by several researchers. However, I maintain that the model of generation with the EOS as defined here, is independent of semantic representation so long as the procedures that consult the EOS can provide the required information to the generation grammar and so long as the required information is contained in or deducible from the semantic representation. We shall see how this is done in the section on consulting the EOS.

In Section 3.2 I discuss XPs and the basis in X-bar theory<sup>48</sup> of the formalism proposed, while Section 3.3 deals with the transformational component of GB theory. Related work in logic grammars is discussed in Section 3.5, and Section 3.6 summarises the formalism presented here and mentions some areas for future work.

---

<sup>48</sup>X-bar theory deals with the the subcategories of an XP. Thus, X-bar is a subcategory of XP and X is a subcategory of X bar.

### 3.2. The role of X-Bar Theory

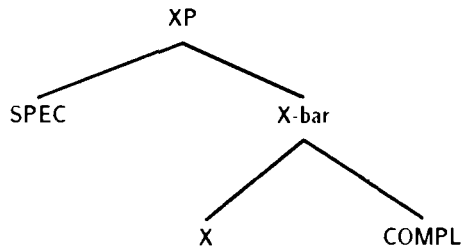
The basis of current X-Bar Theory is that the phrase structure component of the Base Grammar is given by the rules proposed by [Stowell 81] and since adopted in [Chomsky 86]. I shall assume that the language specific parameters which determine word order are correctly set for Spanish. (These parameters determine, for example, which of the rules (3.1.a) and which of the rules (3.1.b) apply to a particular language.)

(3.1.a)  $XP \rightarrow SPEC\ X\text{-bar}$  or  $XP \rightarrow X\text{-bar}\ SPEC$

(3.1.b)  $X\text{-bar} \rightarrow X\ COMPL$  or  $X\text{-bar} \rightarrow COMPL\ X$

Viewed graphically, this means that the phrase structure of the Base<sup>49</sup> component of the grammar has the fundamental structural form<sup>50</sup>:

(3.2)



In this notation, **X** stands for the one of the clausal categories **C** or **I** (for Complementiser or Inflection respectively), or for one of the lexical categories **V**, **N**, **A**, or **P** (for Verb, Noun, Adjective, or Preposition). **X-bar** is called a *projection of X*, and

---

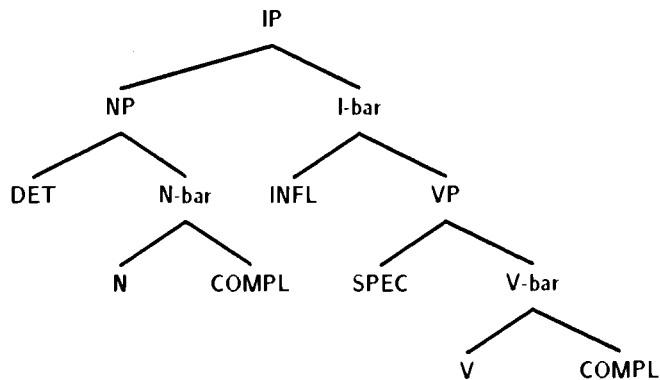
<sup>49</sup>The Base component of a transformational grammar is that part which generates the initial structure called D-structure. The Transformational component subsequently changes D-structure into another form called S-structure.

<sup>50</sup>Here and throughout this dissertation I ignore the modifier constructions which adjoin to these structures in the Base.

**XP** is called the *maximal projection of X*<sup>51</sup>.

It is inherent in X-Bar theory that each of these structural units introduces a terminal **X** which may or may not have phonological or orthographical content. The specifier (**SPEC**) position and the complement (**COMPL**) position may be filled with another of these **XP** structures in turn. For example:

(3.3)



Of course the **SPEC** may be empty or a lexical item (such as a determiner, when the **XP** is an **NP**) and either a **SPEC** or a **COMPL** may be absent.

In a very much simplified<sup>52</sup> view of the logic grammar formalism developed here, the X-bar rules (3.1) are replaced by rules of the general form (3.4). In section 3.3 we shall see how consultation with the EOS is added:

<sup>51</sup>See [Radford 81] [Chomsky 82a] and [Stowell 81] for a more complete discussion of X-bar theory.

<sup>52</sup>These are clearly context free rules, and lack any mechanism for actually building the parse tree, and for consulting the context of generation.

(3.4)

$$\text{xp}(\text{Cat}, \text{PT}, \text{Graph}, \text{Inf}) \rightarrow \text{spec}(\text{Cat}, \text{PT0}, \text{Graph}, \text{Inf0}), \text{xbar}(\text{Cat}, \text{PT1}, \text{Graph}, \text{Inf1}).$$

$$\text{xbar}(\text{Cat}, \text{PT}, \text{Graph}, \text{Inf}) \rightarrow \text{x}(\text{Cat}, \text{PT0}, \text{Graph}, \text{Inf0}), \text{compl}(\text{Cat}, \text{PT1}, \text{Graph}, \text{Inf1}).$$

where:

**Cat** stands for the syntactic category.

**PT** is an internal parse-tree structure being built up out of **PT1** and **PT0** from the subcategories.

**Graph** is the conceptual graph from which the sentence is being generated.

**Inf** is the informant<sup>53</sup> list of features; some being passed on to subsequent rules, some to be unified with unbound informants.

### 3.3. The Transformational Component of GB Theory

GB theory has a Base component, a Transformational component, and several sets of principles [Chomsky 82a]. Here we shall be concerned only with the transformational component which allows XPs and Xs to move from their original position in the syntactic structure either to empty positions (e) or to adjoin to XP positions. It is due to this component that we can achieve a clean separation of syntax and semantics because transformations permit the generation of more than one sentence with the same meaning. The following is assumed: The Base component generates the D-structure(s) associated with<sup>54</sup> a given semantic representation. The Transformational component then operates on the D-structure to produce S-structures with the same meaning. More than one S-structure may result.

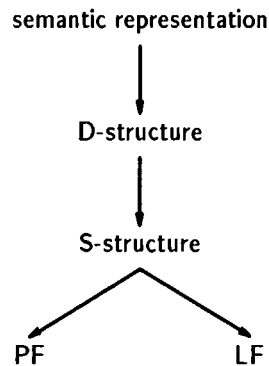
Many NL phenomena have been studied that express the same meaning in more than

---

<sup>53</sup>I shall use the term *informant list* for the argument of the Prolog Predicate containing the list of syntactic features.

<sup>54</sup>There may be more than one D-structure associated with a single semantic representation.

## (3.5)



one way syntactically. Transformations are one way of expressing the syntax of such sentences. The motivation is that a grammar that has to deal with phenomena as diverse as active and passive sentences, topicalised or extraposed NPs, weak crossover effects, Wh-movement and weak subadjacency violations must endeavour to achieve great expressive power. Failure to do so has extremely severe consequences: namely a proliferation of rules to deal with each phenomenon individually, and a consequently large set of constraints to prevent interference of the sets of phenomenon-particular rules with each other. I feel that it is exactly the difficulty in achieving expressive power that has led to computational models of analysis and generation that exhibit the following undesirable features:

- A proliferation of rules to deal with the above-mentioned phenomena individually.
- Difficulty in demonstrating, for large fragments of NL, that sentences subject to the above phenomena can have the same semantic representation.
- Difficulty in incorporating semantic structures in the grammar at the implementation level.

The result has been severely limited grammatical coverage. The corpus of related computational research will be discussed in section 5.

It is not the purpose of this Chapter to argue extensively in favour of transformational



grammars. However, the tendency, computationally, has been to shy away from this model of generation. The reason is usually based on arguments of complexity, incomprehensibility, incompleteness, and lack of semantics. I would make two points here: One, that grammatical coverage is lacking in implementations in part because they lack a transformational component. Two, that I believe the present state of GB theory [Chomsky 86] provides a window for semantics within the environment of syntax (EOS), (and context of discourse or LF').

I suggest that a transformational component, Case theory, Theta-theory, and all the other principles of GB theory that have evolved (painfully) over the last two decades should be carefully considered. They are supported by vast amounts of scholarship and data. We should not shy away from the apparent complexity of a model that attempts to achieve explanatory adequacy. Instead, we should attempt to embrace the goals of expressive power and explanatory adequacy, and in so doing, attempt to achieve the computational elegance we all seek<sup>55</sup>

We have argued elsewhere that Discontinuous Grammars [Dahl 84] [Dahl and Abramson 84] [Dahl, Brown et al 86] and, in particular, Constrained Discontinuous Grammars [Dahl and St-Dizier 86] [Dahl and Brown 86] provide the kind of expressive power for transformations that is needed by writers of logic grammars for NL. In such grammars, transformations are easily expressed.

---

<sup>55</sup>An argument can be made [Victor Castel, private communication] that in GB the phrase structure rules are indeed less numerous but the complexity is then carried over to all other principles and filters. The reply is that the principles of GB are much more general than the rule justification needed in non-transformational models.

### 3.4. Topicalisation - an example of a transformation

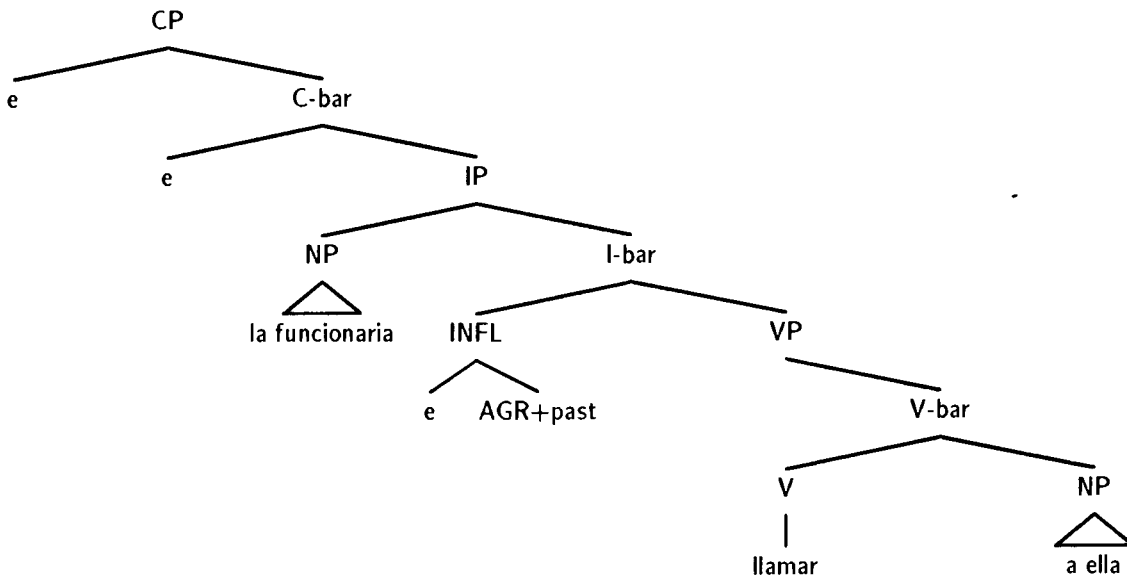
Consider the sentence (3.6) where (3.6.a) and (3.6.b) are assumed to have the same meaning with emphasis on the NP [a ella].

(3.6.a) **La funcionaria**  $la_i$  llamaba a  $ella_i$ .  
*The clerk called HER.*

(3.6.b) **[A ella]**  $i$ ; **la funcionaria**  $la_i$  llamaba.  
*HER, the clerk called.*

It is assumed that both sentences have the same D-structure (3.7), and that this D-structure is generated only once.

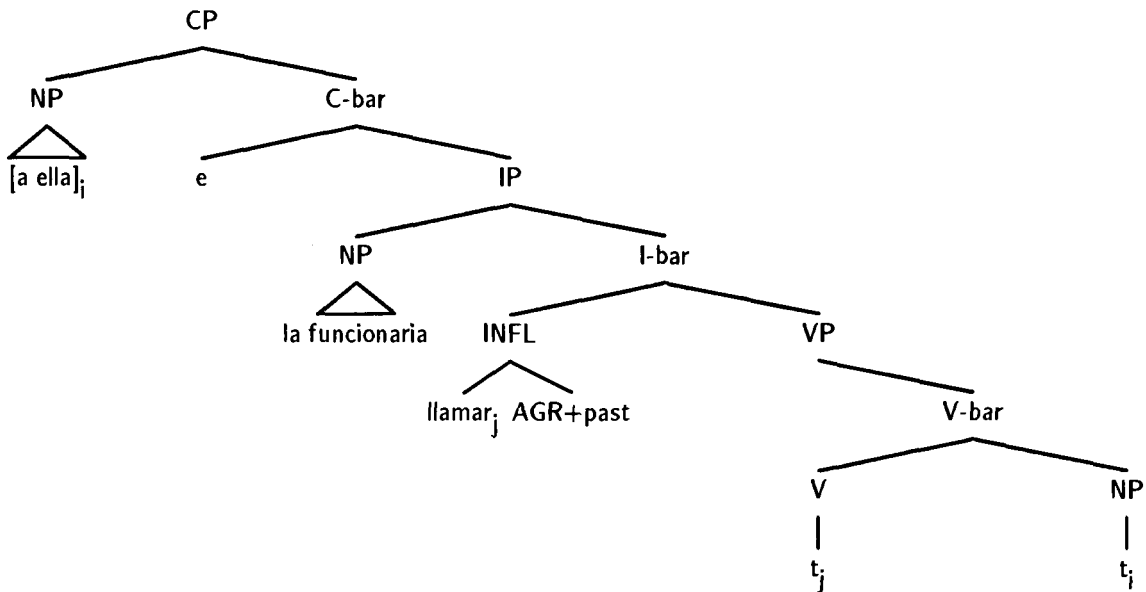
(3.7) **La funcionaria**  $la_i$  llamaba a  $ella_i$ . (D-structure)  
*The clerk called HER.*



If no topicalisation transformation occurs, sentence (3.6.a) results<sup>56</sup>. If the topicalisation transformation occurs, moving the NP [a ella] to the front of the sentence, sentence (3.6.b) results with the structure of (3.8).

<sup>56</sup>In the case of (3.6.a) emphasis would have to be obtained phonologically.

(3.8) [A ella]<sub>i</sub> la funcionaria la<sub>i</sub> llamaba. (S-structure)  
 HER, the clerk called



### 3.5. Transformations and semantics

If we make the hypothesis that the semantic representation needs to be consulted only in the generation of D-structure by the Base component of the grammar, and not by any further (transformational) component, we have the basis for a tidy separation of syntax and semantics. The questions of how such a consultation can be achieved and on what basis in the grammar remain to be answered.

It is the thesis of this Chapter that such a consultation should be made at the XP level of the grammar. Each and every time an XP is to be generated, a consultation with the semantic representation should take place. The intuition is that an XP represents, generally, one lexical substitution in the abstract phrase structure of the sentence. Further, it is the semantic constraints on individual words that often need to be examined. To this end, I introduce the notion of Environment of Syntax (EOS) which contains the following components, the uses of which are examined in the next section:

- An augmented lexicon (a local lexical context).
- A categorical type (a local syntactic context).
- A semantic representation for one sentence (a local semantic context).
- A semantic-type hierarchy (a local pragmatic context).
- An R-expression 'stack' (a local discourse and sentential context).

Now we have a means of separating the rules of the grammar from the semantic representation itself. The semantic representation or its decompositions should not appear in the grammar rules. Instead, the Base component must be provided with a means of consulting the EOS. The transformational component may then provide alternative syntactic expressions of the meaning.

Such a strategy is advantageous if a means exists of expressing and automatically generating such transformations. The Discontinuous Grammar (DG) formalism discussed at length in Chapter 6 provides rules that do exactly that. To give the flavour of such rules, consider once more the topicalisation example above.

In this case an NP [a ella] with Case and  $\Theta$ -role assigned is free to move to an empty XP position<sup>57</sup>. The DG rule (3.9) expresses such a transformation. In this rule, the XP generated by the predicate `xp(np,PT1,G,[no_theta_role,no_case])` would normally become empty (e) if the rule were not applied. Similarly, the XP generated by the predicate `xp(nptrace,PT2,G,[TR,Case])` becomes a trace (t) in the internal parse-tree representations, when (3.9) is applied.

---

<sup>57</sup>Subject to the constraints of Subjacency and Government.

(3.9)  
`xp(np,PT1,G,[no_theta_role,no_case]), skip, xp(np,PT2,G,[TR,Case])`  
`-->`  
`xp(np,PT1,G,[TR,Case]), skip, xp(nptrace,PT2,G,[TR,Case])`

where:

`np` is the syntactic category.

`PT1` and `PT2` are internal parse-tree representations.

`G` is the semantic representation, and

`[TR,no_case, . . . ]` etc. are the informant lists of features (mentioned in (4)) that represent  $\Theta$ -role and Case respectively.

Discontinuous Grammar rules allow the grammar writer both to express transformations, and to achieve a clear separation of the semantic representation and the grammar rules. The Base generation rules make use of the EOS and generate a D-structure which is then transformed by the DG rules. We can now turn our attention to the EOS and the separation of syntax and semantics.

### 3.6. A model of environment of syntax for NL generation

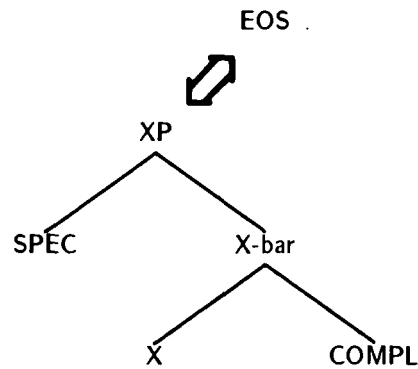
It is the fundamental hypothesis of my model that each XP introduces one consultation with the environment of syntax (EOS) of generation, shown as the Prolog call `consultcontext` in the following skeleton of a rule (3.10). It is through this consultation that syntactic features are established or verified:

(3.10)  
`xp(Cat, . . . , Graph, . . . ) -->`  
`{consultcontext(Cat,Graph, . . . )},`  
`spec(Cat, . . . ),`  
`xbar(Cat, . . . ).`

This may be viewed graphically as (3.11):

Notice that in (3.10) the consultation with EOS has arguments for both the category

(3.11)



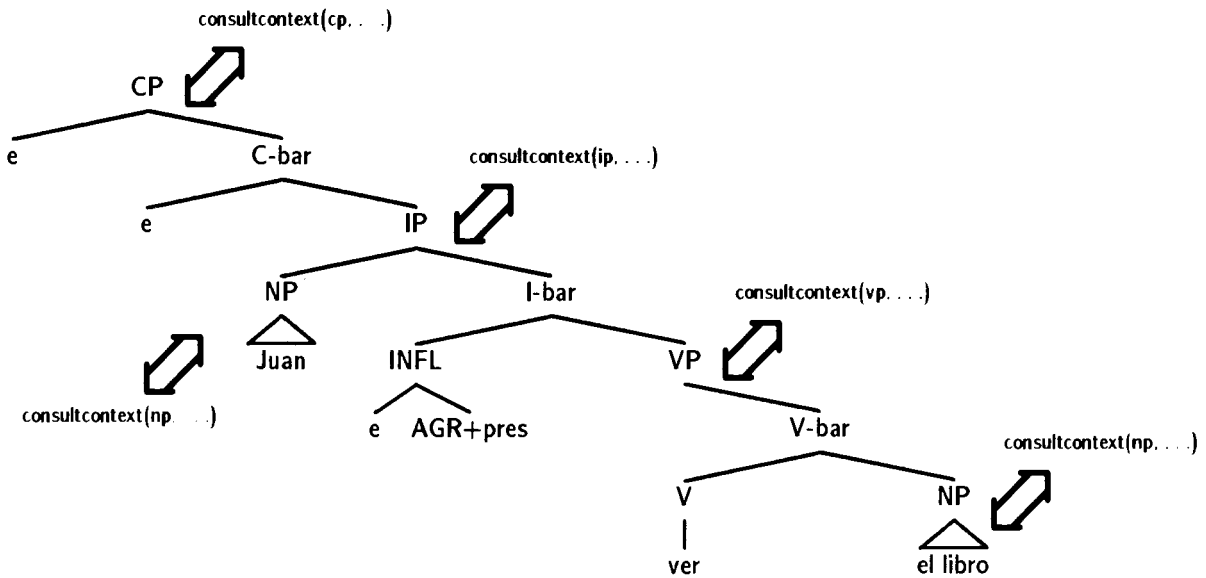
(**Cat**) of the **XP** and the semantic structure, in this case the conceptual graph (**Graph**). This allows different procedures in the context consultation process based upon the syntactic category.

I shall examine the **consultcontext(Cat, . . .)** procedures for the categories **CP, IP, VP, NP, PP** in following sections. First, however, let's take a brief look at the lexicon and the information encoded in the verb and noun entries. Then the **consultcontext(np, . . .)** is considered to show how contextual constraints may be inserted. The examination of the **consultcontext(vp, . . .)** procedure shows how the combination of transformations and an augmented lexicon allow the same conceptual graph to be used in the generation of both active and passive sentences.

### 3.6.1. The components of context

Components of this model of EOS are an augmented lexicon, a semantic representation of the sentence to be generated, a stack of R-expressions, and a semantic-type hierarchy. There is also a set of procedures, called **consultcontext**, for consulting the EOS. Such procedures include rules for semantic-type checking, determination of animate/inanimate status of R-expression, setting dialectal parameters, and determining whether an R-

(3.12)



expression can be pronominalised, in addition to lexical insertion and determination/verification of  $\Theta$ -roles.

In order to provide all the above, the lexicon must be augmented with subcategorisation, Case, and  $\Theta$ -role/type information. Some examples taken from the lexicon of Spanish will illustrate:

(3.13)

*noun(Token, Word, SubCategorisation, Agreement, ThetaType, Animate).*

**noun(petrol, bencina, [cp], agr(3, s, f), liquid, inani).**

This entry in the lexicon states that the semantic token **petrol** is associated with the lexical item **bencina** which may subcategorise with a **cp**. The agreement features of this lexical item are **third person, singular, feminine**, its  $\Theta$ -type is **liquid**, and it is **inanimate**.

## (3.14)

*verb(Token,Word,Subcategorisation,List-of-argument-features).*

**verb(buy1,comprar, np, [[agt,nom,Def1|Rest1],  
[theme,acc,Def2|Rest2]]).**

**verb(buy1,comprar, [np,pp(por,np)], [[no-tr,nom,Def1|Rest1],  
[theme,nocase,Def2|Rest2],  
[agt,obl,Def3|Rest3]]).**

The verb entries shown are for the verb whose semantic token is **buy1** associated with the lexical entry **comprar** in Spanish. The first entry corresponds to active usage, where the subcategorisation is an **np** which is assigned the  $\Theta$ -role [**theme**] and accusative Case [**acc**]. The second entry corresponds to the passive usage of the verb. Here the first subcategorisation category is **np** which is assigned  $\Theta$ -role [**theme**] but to which no Case is assigned (indicated with feature [**nocase**]). The second subcategorisation category is **pp(por,np)** indicating that a PP with the preposition **por** must be generated, whose case is [**obl**]<sup>58</sup> and whose  $\Theta$ -role is [**agt**] (for **agent**). Notice that the second entry for **buy1** has an *extra* argument. Now let's leave the example and return to the discussion of the EOS.

The semantic representation for a sentence in the EOS need only provide the logical argument status of the predications. In this model, a heavy burden is placed on the procedures that consult the semantic representation through the EOS. These procedures must be able to return or verify  $\Theta$ -roles. The  $\Theta$ -roles may or may not be encoded directly in the semantic representation. Prelexical processing (such as the levels of Grammatical Relations discussed in [Permuter and Rosen 84], the lexical functions of Melcuk used in [Boyer and Lapalme 85], or the Structure Building Component of

---

<sup>58</sup>Oblique case.



[DeArmond 87]) may alter (or follow changes to) the  $\Theta$ -roles found in the augmented lexicon

The stack of R-expressions is used for reference and pronominalisation. This stack is consulted when a noun is being inserted. The process consulting the stack must decide whether an R-expression may be pronominalised or not. At present, the generation grammar allows both pronominalised and full NPs to be generated through the non-deterministic proof procedure of Prolog. Some dialectal processing is possible at this level. For example whether or not to double direct object clitics is a dialectal feature in Spanish, and is fixed at the level of IP when the EOS is consulted. A type hierarchy and semantic type processing are common constraints on lexical insertion [Dahl 77]. For example the agent of verbs such as see must be animate

### 3.6.2. Consulting context from NP

When an NP consults the EOS, the object of this call is to retrieve the lexical head of the NP, and any determiner, from the EOS, and to either find a  $\Theta$ -Role, or to verify that the  $\Theta$ -Role is correct. Additional information such as agreement information, and whether, for example, the concept is an animate or inanimate one should also be retrieved. An additional task of the call (other than to verify or retrieve information) is to check constraints on the concept within the EOS.

These tasks have proved to be an adequate treatment of EOS for the purposes of NL generation in a proposed prototype NL translation system [Dahl, Brown et al 86]. The specific call for consultation with the EOS is given in

(3.15)

```

xp(np,PT,Graph,[TR,Case,Def,Agr,Animate,Pronominal]) -->
  {consultcontext(np,Word,Graph,SubCat,[Def,Agr,TR,Animate,Pronominal])},
  spec(np,PT1,[Def,Agr]),
  xbar(np,PT2,Graph,Word,SubCat),
  {PT=..[np,[TR,Case,Def,Agr],PT1,PT2]}.

```

```

consultcontext(np,Word,Graph,SubCat,[Def,Agr,TR,Animate,Overt]) :-
  processgraph(Graph,(Def:Token),TR),
  noun(Token,Word,SubCat,Agr,ThetaType,Animate),
  checkexternal(np,(Def:Token),Overt),
  constraint(np,TR,ThetaType),
  checktype(ThetaType,Animate).

```

where

**np** is the syntactic category of the XP.  
**Graph** is the semantic representation of the sentence.  
**[TR,Case,Def,Agr . . . ]** is a list of syntactic features.  
**TR** is the  $\Theta$ -role of the NP.  
**Case** is the morphosyntactic Case (nominative, dative, etc.).  
**Def** is the definiteness feature (definite, indefinite) of the NP  
**Animate** is the animate/inanimate status feature.  
**Pronominal** is the pronominalisation feature.  
**SubCat** is the subcategorisation category of the noun's complement  
**Word** is the lexical item (the noun).  
**ThetaType** is the semantic-type of the noun.  
**PT, PT1, PT2** are internal parse-tree structures.

Looking at this process in more detail, the first call to **processgraph**, inputs the name of the semantic structure, (**Graph**), and outputs a token, (**Token**) for a lexical item plus a definiteness marker, (**Def**), indicating definite, indefinite, or no marking. The  $\Theta$ -Role [**TR**], may be *input in* which case it is verified, or it *may be unbound in* which case it must be returned bound to some  $\Theta$ -role.

(3.16)

```

processgraph(Graph,(Def:Token),TR),

```

It is most important here to note that the structure of the semantic representation is not specified. It is sufficient to specify what the action of **processgraph** is. There is no requirement for a specific type of semantic structure. Neither is there a specification of

how the  $\Theta$ -Role is to be found in the semantic structure. The role of **processgraph** is to return the **Def** marker and to verify or return the  $\Theta$ -roles. The  $\Theta$ -roles and **Def** may not be expressed specifically in the semantic representation, but may instead be inferred.

The next task is to consult the lexicon using the semantic token to find a corresponding lexical item (There may be more than one [Brown and Sempere 87]).

(3.17) **noun(Token,Word,SubCat,Agr,ThetaType),**

This call will return the actual lexical item (**Word**) and also provide agreement(**Agr**), semantic type (**ThetaType**), and subcategorisation frame (**SubCat**)<sup>59</sup> information

The above two calls are essential in this model. Now we can begin to check constraints on the generation and to make effective use of the context. First a check to see whether the noun can be pronominalised by checking the stack of previously named concepts. **Overt** is the output informant and is set as **+overt** or **-overt** by this call:

(3.18) **checkexternal(np,(Def:Token),Agr,Overt),**

Next a check that the noun's **ThetaType** is actually consistent with the required  $\Theta$ -Role **TR** by checking the semantic hierarchy:

(3.19) **constraint(np,TR,ThetaType),**

Then a determination of the animate/inanimate status of the concept.

(3.20) **checktype(ThetaType,Animate).**

---

<sup>59</sup>The subcategorisation frame gives the category of expected complements (COMPL) of the XP

### 3.6.3. Consulting context from VP

When a VP consults the EOS, it knows whether it is generating an active or a passive sentence through a **Voice** informant. This allows pre-lexical processing specific to each voice. The setting of **Voice** to active [**actv**] or passive [**pass**] may be done non-deterministically by the grammar, or it may be specified in the call to generate a sentence (from the CP level). If the feature **Voice** is unbound in the call to generate a CP, the grammar will first set the feature to [**actv**] and try to generate an active sentence, then backtrack and set the feature to [**pass**] and try to generate a passive sentence.

(3.21)

```

xp(vp,PT,G,[TNS,Voice,SubjInf,ComplInf]) -->
  {consultcontext(vp,W,G,SubCat,Voice,-bnft,[SubjInf,ComplInf]),
  spec(vp,PT1,G,[Voice]),
  xbar(vp,PT2,G,W,SubCat,ComplInf),
  {PT = [vp,[TNS],PT1,PT2]}.

```

where.

- vp** is the syntactic category of the XP
- G** the semantic representation of the sentence.
- [TNS,Voice,SubjInf,ComplInf]** a list of syntactic features.
- TNS** the tense feature (finite or infinite).
- Voice** the voice feature (active or passive).
- SubjInf** the list of informants of the subject.
- ComplInf** the list of informants of the complement.
- W** the lexical item (here a verb).
- SubCat** the subcategorisation category of the complement.
- PT PT1, PT2** are internal parse-tree structures.

The first job of **consultcontext** here is to find the predication<sup>60</sup> in the semantic representation and to find the  $\Theta$ -roles of its arguments. Next is to find a corresponding verb in the lexicon and verify that the verb's lexical entry can support these  $\Theta$ -roles, either in the active or the passive voice.

---

<sup>60</sup>It is assumed that the predication is unambiguous.

The first rule (3.22) for `consultcontext(vp, . . .)` deals with the active voice. In this case, two lists represented above as `SubjInf` and `ObjInf`, are returned with the features of the external and internal arguments of the lexical entry found by `verb` corresponding to the  $\Theta$ -roles found by `processgraph`. Recall that the `Token` for this verb must unify with `Token` returned by `processgraph`, and notice that the  $\Theta$ -roles returned by `processgraph` must unify with the first elements of the feature lists.

(3.22)

```
consultcontext(vp,Word,Graph,SubCat,actv,[[SubjTR|Rest1],
                                           [ComplTR|Rest2]]) :-
    processgraph(Graph,Token,[SubjTR,ComplTR]),
    verb(Token,Word,SubCat,[[SubjTR|Rest1],[ComplTR|Rest2]]).
```

The second rule (3.23) deals with the passive<sup>61</sup> voice. Here we have an example of *pre-lexical processing* corresponding in nature to the work discussed in [Permutter and Postal 83]. The call to `processgraph` returns two  $\Theta$ -roles, `AgtTR` and `ThemeTR`, while the lexical entry found by `verb` has three argument feature lists. Thus, the lexical entry for the verb found by the rule for active voice above will not succeed for this rule for passives. A second entry in the lexicon is needed for the passive. By allowing, or not, such a lexical entry, the composer of the lexicon can control the passivisation of a verb. Of course, careful attention must be paid to the  $\Theta$ -roles. As is well known, some verbs passivise, some do not.

(3.23)

```
consultcontext(vp,Word,Graph,SubCat,pass,[[no-tr|Rest1],
                                           [ThemeTR|Rest2]],
```

---

<sup>61</sup>Recall that the decision here has been to control the passivisation of a verb through the presence or absence of suitable entries in the lexicon. Other methods of control are, of course, possible which might use a single lexical entry plus a flag in the lexical entry to indicate the possibility, or not, of passivisation. If, in fact, the passivisation of Spanish verbs is predicatable (R. DeArmond, personal communication), then neither flag nor separate lexicon entry is necessary. A procedure could, for example, examine whether the verb is transitive and does not assign `theme` to its external argument. If these conditions were met, a 'passive' entry could be created dynamically.

[AgtTR|Rest3]] :-

**processgraph(Graph,Token,[AgtTR,ThemeTR]),**  
**verb(Token,Word,SubCat,[[SubjTR|Rest1],[ThemeTR,Rest2],[AgtTR|Rest3]].**

A glance back to the lexical entries (3.14) given as examples earlier will illustrate the difference in concrete terms. With these entries, the following active/passive pair of sentences have been generated<sup>62</sup>:

(3.24.a) **El médico compró la bencina.**  
*The doctor bought the petrol.*

(3.24.b) **La bencina fue comprada por el médico.**  
*The petrol was bought by the doctor.*

### 3.6.4. Consulting context from PP

The process of consulting the EOS from PP is similar to that for VP for those prepositions that assign  $\Theta$ -roles. I have shown elsewhere [Brown, Pattabhiraman et al 87] how this applies to the generation of small clauses of the PP type. As with verbs, the preposition may be thought of as having an external and an internal argument. The lexical entry for the preposition must therefore contain  $\Theta$ -role and Case information for these arguments, as well as a subcategorisation category for the internal argument.

### 3.6.5. Consulting context from CP

At the CP level of syntax the EOS must provide basic information about the clausal level of the sentence being generated. The informant **Level** determines whether the clause is a matrix (**mtrx**) clause, a subordinate clause (**sub**), or a relative clause (**rel**). The other basic clausal parameters set at this level are **Voice** and **Type**. **Voice** may be active (**actv**) or passive (**pass**). **Type** indicates whether a matrix sentence is declarative (**dcl**).

---

<sup>62</sup> recognise that such true passives are rare in Spanish.

imperative (**impr**) or a question (**que**) Additionally a  $\Theta$ -role is available when the CP is a subordinate clause

(3.25)

```
consultcontext(cp, __, Graph, SubCat, Informants) :-
    SubCat=ip,
    Informants=[Level, Voice, Type, ThetaRole].
```

The subcategorisation of CP is always taken to be IP. Matrix clause rules for **consultcontext** that provide (nondeterministically) both active and passive voice for declarative sentences are the following. The  $\Theta$ -role for a matrix clause is assumed to be null.

(3.26)

```
consultcontext(cp, __, Graph, SubCat, Informants) :-
    SubCat=ip,
    Informants=[mtrx, actv, dcl, null].
```

```
consultcontext(cp, __, Graph, SubCat, Informants) :-
    SubCat=ip,
    Informants=[mtrx, pass, dcl, null].
```

### 3.6.6 Consulting context from IP

For IP the **consultcontext** call must determine whether the clause is tensed (+**tns**) or not (-**tns**) and unify the **TNS** informant with the appropriate value. This informant is passed on to the complement subcategory, VP. The specific tense (**Tense**), if any is also found for subsequent use in **INFL**.

(3.27)

```
consultcontext(ip, __, Graph, SubCat, [TNS, Case, Overt, KOvert, Tense]) :-
    SubCat=vp,
    findialect(Graph, Dialect),
    resolveovert(Dialect, Overt, KOvert, Case),
    findtense(Graph, Tense).
```

Other informants used here are specific to Spanish and involve the **OVRT** feature of the

elements of the of the direct object clitic chain. These are **Case**, **Overt**, and **KOvert**. The use of these informants will be explained in Chapter 5 where the generation of clitic chains and the **resolveovert** call are discussed at length. The associated **Dialect** informant is specific to Spanish and is used to determine whether clitic doubling should occur for direct object clitics. This informant is passed from the **findialect** call to the **resolveovert** call.

### 3.7. Related work in Automatic NL Generation

Several logic grammar formalisms have been proposed that are intended to provide a strict separation of syntax and semantics in order that the writer of a grammar may easily express syntax, semantics, and their interrelation. In this section I shall briefly review them and discuss how they differ from the formalism presented here. In addition, I shall look at other work in automatic NL generation which has been generally limited by the semantic representations chosen, by the context free nature of the generation grammars, and by the interface between grammar and semantic representation.

The Modular Logic Grammar of [McCord 85] is a good example of the separation of syntax and semantics in a logic grammar formalism. I fail to see the linguistic justification for the *strong* and *weak* terminals introduced in this work. The motivation seems to have been to find a way to separate the semantic structure from the phrase structure rules of the grammar. However, this approach falls into the trap of actually including the semantic structures in the compiled Prolog clauses. (Of course this is a viable option for a simple and constrained syntax.) Contextual processing is also precluded in such a model. I suspect also that the semantic structures capable of being used in such a system are similar to those provided by the Metamorphosis Grammar [Colmerauer 78] based systems characteristic of earlier work such as [Dahl 77] or [McCord 81].



The work of [Boyer and Lapalme 85] based on associating text with meaning through unification uses the interesting Meaning Text (MT) lexicon of Melcuk. This contains relations (lexical functions) between the words of the lexicon, and may be thought of as containing very specific subcategorisation information about the lexical entries. A semantic component of the lexicon contains semantic graphs, semantic nets with labelled directed arcs associated with fragments of syntax trees, and with conditions on the nodes of the trees. Initial syntactic structures are formed from the fragments. Transformations modify the syntactic structure before text is generated. I suggest that these transformations, using lexical rules with the MT lexical functions, are one type (perhaps the only type) of pre-lexical-insertion processing in GB theory and as such belong in the **consultcontext** process in my generation formalism, where they would serve the purpose of creating (dynamically) a lexical entry to suit that demanded by the syntax of the grammar.

Passives can be generated by the system of [McKeown 83], based on a determination of focus with respect to the previous sentences generated. In my opinion, this is context of discourse processing and should be directed towards providing a semantic representation for a single sentence with or without an indication of focus. Such a representation would then be interpreted as part of the EOS by my grammar which does not now include mechanisms for focus. For example, McKeown's rule: *If the focus is on the AGENT then generate an active sentence; If the focus is on the OBJECT then generate a passive.* should be part of the EOS processing at the CP level in my formalism, and used to bind the **Voice** informant to **actv** or **pass**. McKeown's system does not appear to offer any transformational component. I do not know if it would handle a topicalised NP as a focused element. The grammar appears to be very simple.

The strategic-composer plus syntactic-composer of [Danlos 85] provides some processing

at what I have called the EOS level. This processing by the syntactic-composer is very language specific, and provides a basic syntactic schema for the final syntax. However, post-D-structure transformations are mixed together with the EOS-level processing, making the linguistic basis of the work uncertain. For example, the schemas for the sentences are active or passive in nature *before* processing by the syntactic-composer. This was done to cut the conceptual representations into sentence sized chunks and to order those chunks in a reasonable way. In doing so, the possibility of expressing a conceptualisation in either active or passive voice is lost.

Definite Clause Translation Grammars (DCTGs) of [Abramson 84] are a subset of MGs that also provide a formal separation of syntax and semantics in a logic grammar formalism. The failing here again is that EOS as such plays no part. The role of the lexicon is not specified, and other constraints at the EOS level play no part in the formalism. In addition, and this is common to most of the formalisms that attempt to separate syntax and semantics, a semantic part is required in each rule of the grammar. In my formalism, the EOS (and thus semantics) plays a part only in the rules for XPs. It may be possible to write a DCTG in such a way as to create the same effect as that found in my system, that is, to provide EOS level processing, rather than simple semantic structure interpretation, because DCTGs allow a complete rule for the semantic processing to be embedded in each DCTG rule (and this rule determines or can help to determine a set of attributes - my features - associated with the phrase structure part of the DCTG rule.) Still not provided is a transformational component.

Some researchers have suggested that their NL analysis programs can be run *backwards* and used as generators. In particular the variants of MGs (for example [Dahl 77] [Giannesini et al 85] ) treat the semantic representation as an informant, or argument,

in the rewrite rules. Such systems have a predicative semantics with no marking of the argument arcs of the predications. The constraint imposed by such hierarchical systems is that the semantic representation *must* be functionally compositional. The interface between syntax and semantics in such systems has a theoretical basis in theories of generalised quantifiers. Only a small fragment of grammatical coverage has been explored in these systems

Programming such a system is difficult especially for transformational constructions. The result has been severely limited grammatical coverage or an *ad hoc* semantics (or a new logic grammar formalism designed for a particular problem such as the MSGs [McCord 82]). Thus in the following example, the informants  $X$ ,  $Y$ , and  $P(X,Y)$ , are the semantic structures of their respective predicates:

- (3.28.a)  $\text{sent}(P(X,Y)) \quad \rightarrow \quad \text{np}(X), \text{vp}(P(X,Y)).$
- (3.28.b)  $\text{vp}(P(X,Y)) \quad \rightarrow \quad \text{verb}(P(X,Y)), \text{np}(Y).$
- (3.28.c)  $\text{np}(X) \quad \rightarrow \quad \text{noun}(X).$
- (3.28.d)  $\text{noun}(\text{Kirsten}) \quad \rightarrow \quad [\text{Kirsten}].$
- (3.28.e)  $\text{noun}(\text{Anders}) \quad \rightarrow \quad [\text{Anders}].$
- (3.28.f)  $\text{verb}(\text{loved}(X,Y)) \quad \rightarrow \quad [\text{loved}].$

This set of rules may be used to generate the sentence

(3.29) **Kirsten loved Anders.**

with the associated semantic structure  $\text{loved}(\text{Kirsten}, \text{Anders})$ . It cannot however generate the sentence

(3.30) **Anders was loved by Kirsten.**

MGs allow transformations through symbol rewriting with rules like (3.31).

(3.31)  $\text{np}(X), \text{vp}(P(X,Y)), \text{np}(Y) \rightarrow \text{np}(Y), \text{aux}, \text{vp}(P(X,Y)), \text{np}(X).$

However, to generate the above sentence with a logic grammar without transformations requires three more rules:

- (3.32.a) **sent(P(X,Y))** --> **np(Y), vp(P(X,Y)).**  
 (3.32.b) **vp(P(X,Y))** --> **aux,verb(P(X,Y)), np(X).**  
 (3.32.c) **aux** --> **[was].**

Clearly rules (3.28.a) and (3.32.a), and (3.28.b) and (3.32.b) will interfere with each other at the semantic level. Using rule (3.28.a) with rule (3.32.b) will lead to semantic nonsense in analysis or syntactic nonsense in generation.

Attempts to fix this problem by adding an extra *active/passive* informant will only pose problems when Case is incorporated into the theory, and in any case requires rule duplication<sup>63</sup>. At the very best, we have an *ad hoc* set of rules to deal with just one of the problems attacked by the more general transformational model.

A generation system based on the traversal of conceptual graphs is reported in [Sowa 84]. In this work a generation grammar *per se* did not exist, but was embedded in the algorithm for traversal. A particular syntactic role was presumed for each relation in the conceptual graph. This system was equivalent to a context free grammar, with all the attendant shortcomings in expressive and explanatory power. Such a system is limited and very language specific. In addition, only limited use may be made of the particular form of lexicon which has no subcategorisation frames but which does have thematic roles encoded as relations.

---

<sup>63</sup> Thus, an NP will have nominative case if it precedes the verb and objective case if it follows the verb. However such *ad hoc* rules are known to lead to problems when applied to extraposed NPs and non-matrix clauses.

### 3.8. Towards a more powerful automatic NL generation system

The very brief summary of related work has pointed out shortcomings in either the expressive power of the grammars, the syntax-semantics interface, the use made of the lexicon, or the kind of semantic representation that may be used.

The formal basis for the work in coordination by [Dahl and Abramson 84] [Dahl 84] using Discontinuous Grammars<sup>64</sup> and the work on constraining Discontinuous Grammars [Dahl and St-Dizier 86] has provided the foundation for the more linguistically motivated Discontinuous Grammars used in the work presented in this thesis. Discontinuous Grammars in their original formulation have proved to be incomplete for NL generation, but held the promise of allowing the specification of any transformation.

The system described in this thesis overcomes all of the problems mentioned in connection with previous work, and has been tested for move XP and move-X transformations, as well as for minor movement rules such as affix-hopping. Its characteristics are:

- The EOS-syntax interface is well-defined and is independent of the semantic representation<sup>65</sup> so long as the required information can be obtained or inferred from the semantic representation.
- It has been possible to establish the interface between syntax and EOS by postulating just one consultation with EOS for each XP. This is made possible because the model of syntax is transformational, leaving the Base component to generate D-structure, and allowing transformations to occur in a manner compatible with Case theory and the principles of GB theory. The generation formalism presented here has all the expressive power of a context sensitive

---

<sup>64</sup>Previously called Gapping Grammars

<sup>65</sup>Although of course, a particular semantic representation, conceptual graphs, was chosen for this work.

grammar I conclude that the transformational generative model is a viable model for computational linguistics.

- An augmented lexicon is required in this formalism to give the correct subcategorisation frame of an XP. The lexicon is also augmented with the  $\Theta$ -role information, and with the Grammatical Function information embodied in Case assignment to syntactic argument positions (A-positions). This allows the grammar to make full use of its transformational powers to generate, for example, both active and passive sentences.
- A set of category-specific procedures, called **consultcontext** in this thesis, are necessary to provide the interface between the EOS and the syntactic rules of the grammar.

The question of whether or not a separate lexical entry should be made for each verb that passivises ignores the highly productive nature of the passive construction in English and some other languages. It is of course possible to mechanise the production of the second lexical entry. This possibility however begs the question of which verbs do passivise. It would require an extra argument in the lexical entry to determine whether or not a verb could passivise. I have chosen the separate lexical entry approach as the most general and the most useful in this application.

The question of prelexical processing is a very important one. Here I have shown how one such process may be used to effectively add an extra argument to the set of arguments presented by a semantic representation. This kind of processing may be generalisable along the lines of the work of [Permutter 83] and [Permutter and Rosen 84]<sup>66</sup>. More work needs to be done in the computational aspects of this theory.

Two areas where work remains to be done on the hypothesis (that *Just one consultation with context needs to be made for each XP*) include, firstly, the nature of

---

<sup>66</sup>Perlmutter has shown how the pre-lexical processing is based on general principles for languages as diverse as Kinyarwanda, Georgian and French.

the consultation for APs and, secondly, the relationship of tense, aspect, voice, and mood (all syntactic features, and very strongly language specific) to temporal information in the contextual model.

## Chapter 4

### Clitics as Inflections<sup>67</sup>

#### 4.1. Introduction

Clitics are found in Romance languages. In this chapter I will justify a treatment of direct and indirect object Spanish clitics as verb inflections. Two important questions must be answered here:

- *Where exactly are clitics generated in the Base?*
- *What are clitics, pronouns or inflections?*

#### 4.2. The Nature of Clitics

We begin with a brief examination of the nature of clitics in order to make clear that clitics do not behave like pronouns<sup>68</sup>.

---

<sup>67</sup>This Chapter reviews and extends the essential points covered in the paper *On Clitics as Inflections* by myself and Juan Antonio Sempere-Martinez [Brown and Sempere 85] given at the Second Annual Conference of the Northwest Linguistics Club, at UBC, 22-23 February, 1986 and since translated into Spanish and accepted for publication in the *Revista Argentina de Lingüística* [Brown and Sempere 87]. In a forthcoming paper we discuss the possibility of extending the treatment to all Romance languages

<sup>68</sup>This has been noted by [Strozer 76]. We extend and strengthen the arguments.



### 4.2.1. Separable and Inseparable Affixes

The word inflection in linguistics usually means an affix which is a permanent part of a word. Thus, **stem + inflection** together form an inseparable unit. The word clitic on the other hand applies to elements that do not affect the lexical meaning of **stem + inflection**, but are however affixes to the word. Many languages have such affixes. One difference between traditional inflections and clitics is that clitics may move around. I shall discuss this movement of clitics in what follows. The key point here is that words may have two kinds<sup>69</sup> of affixes: those which are fixed, and those which may move. In what follows, we refer to both kinds of affixes to Spanish verbs as inflections. However, the fundamental difference between the two types should be kept in mind.

### 4.2.2. Proximity to the verb

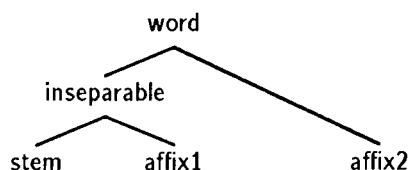
Clitics occur in Spanish as direct or indirect object clitics, as reflexive clitics, and as impersonal clitics. In all cases they are in anteposition or post-position to the verb with which they are associated thematically, with the exception of those clitics involved in

---

<sup>69</sup>Consider the following example from Polish (from R. deArmond, personal communication). A Polish verb may have an inseparable inflection and a separable part:

$$\begin{array}{rcl} \text{nieslam} & = & \text{nies} + \text{la} + \text{m} \\ & & \text{stem} + \text{past.fem.participle} + \text{clitic} \end{array}$$

In this case, the clitic may move to another position in the sentence, not merely to a prefix position. Thus, there seems to be a structure to such words which may indicate moveable and permanent affixation. We could hypothesise the following structure for such words:



*promotion*<sup>70</sup>

Some examples of these clitic roles are given in (4.1): a direct object clitic **lo** (4.1.a), an indirect object clitic **le** (4.1.b), a reflexive clitic **se** (4.1.c), an impersonal clitic **se** (4.1.d), and benefactive dative **me** (4.1.e).

- (4.1.a) **Isabel lo vio a Juan en la calle.**  
*Isabel saw Juan in the road.*
- (4.1. b) **Ella le dio un libro.**  
*She gave him a book.*
- (4.1.c) **No se había afeitado desde hacía una semana.**  
*He had not shaved for a week.*
- (4.1.d) **Le dijo, "¿no se venden navajas aquí?"**  
*She said. "Don't they sell razors here?"*
- (4.1.e) **"Se me escapó la navaja," dijo.**  
*"My razor got lost on me," he said.*

In (4.1.e) we see that clitics may occur in pairs<sup>71</sup>. The clitics associated with the thematic arguments of a verb always occur together and always either in preposition or post-position to the verb.

No other syntactic categories may encroach between clitic(s) and verb, not even a negation (4.1.d). The single exception is the inflected form of the auxiliary in compound tenses (4.1.c). We discuss this phenomenon later.

When a clitic occurs as an argument of an infinitive it is also found adjacent to the

---

<sup>70</sup>We will present a justification for this exception in terms of movement rules for clitics later in this note.

<sup>71</sup>Occasionally three clitics will occur with a verb. It is possible (though rare) for four clitics to occur with a verb. We note these cases are normally ambiguous and provoke different opinions of grammaticality from native speakers. The grammaticality is not a question here. In all cases the group of clitics associated with the thematic arguments of a verb are always found together.

verb In this case, the clitic is always in post-position<sup>72</sup> (4.2.a), except when promotion occurs (4.2.b).

- (4.2.a) **Yo espero visitarla.**  
*I hope to visit her.*
- (4.2.b) **Yo la espero visitar.**  
*I hope to visit her.*

It is interesting to note that the adjacency of clitics to their verbs is enforced even in mixed languages, that is languages that exhibit intrasentential code-switching<sup>73</sup>. In (4.3) for example we see that the clitic **lo** must be associated absolutely with the verb **vio** from its original language, and not with its equivalent **saw** in another language. This is a very strong argument supporting the clitic as an inflection of the verb.

- (4.3.a) **Ella lo vio in the street.**
- (4.3.b) \***Ella lo saw in the street.**
- (4.3.c) **She saw him en la calle.**
- (4.3.d) **Ella saw him en la calle.**

#### 4.2.3. Absolute order of clitics

It has been noted that there is an absolute order of Spanish clitics [Perlmutter 71] [Strozer 76]. The basic order suggested by [Perlmutter 71] is:

se 2nd-person 1st-person 3rd-person

Some refinement of this ordering was provided by [Strozer 76]. This absolute ordering

---

<sup>72</sup>We will use the terms *proclitic* for a clitic which occurs in preposition (anteposition), and *enclitic* for a clitic which occurs in post-position.

<sup>73</sup>Currently being investigated by Juan Antonio Sempere at University of California, Berkeley. Concerning the nature of clitics from a perspective of language transference from language L1 to language L2. Meisel notes that *Clitics have to be considered as belonging to bound morphology; this is confirmed by their syntactic behaviour in general, e.g. no non-clitic constituent can go between the verb and its clitic.* [Meisel 83], page 33. Meisel bases his claim on Poplack [Poplack 80] and on Pfaff [Pfaff 79], in research on code switching constraints.

reinforces the evidence that clitics do not act like pronouns. It is a fact that the order of pronouns is not absolute, as shown in the examples in (4.4).

- (4.4.a) **Se me perdió el libro.**  
*My book got lost on me.*
- (4.4.b) **\*Me se perdió el libro.**
- (4.4.c) **Yo la vi a ella.**  
*I saw HER.*
- (4.4.d) **A ella yo la vi.**  
*Her, I saw.*

#### 4.2.4 Clitics are not NPs

There are many differences between NPs and clitics<sup>74</sup> The following is a brief summary of these differences

- Clitics and personal pronouns are different morphologically.
- NP forms do not distinguish case while (some) clitics do;
- NPs take contrastive or emphatic stress, clitics never do;
- No syntactic category may intervene between clitic and verb while the presence of other categories between NP and verb is commonplace;
- It is not possible phonologically to have a clitic without a verb;
- Clitics but not NPs occur in a fixed surface order;
- NPs but not clitics may take modifiers;
- Clitics and NPs do not occur in corresponding positions in the VP, clitics do not occur in NP positions;
- NPs may be conjoined, clitics may not;
- The topicalisation transformation does not apply to clitics, only to NPs;
- The clitic climbing transformation applies only to clitics, not to NPs;
- Clitics may double NPs, but NPs may not double NPs.

We now look at a representative sample of these arguments in detail in order to emphasise their strength.

---

<sup>74</sup>We accept Strozer's [Strozer 76] (pps 106-113) analysis that clitics are not NPs. Most of the differences mentioned here are hers. Some of her points are omitted because they are weak or incorrect. For example, she claims that clitics may refer to animate or inanimate things while pronouns may refer only to animate ones. This is certainly not true in Spanish. (viz. *Este es el escritorio. Metí los papeles en él.*)

No syntactic category may intervene between clitic and verb while the presence of other categories between NP and verb is commonplace, as demonstrated in sentences (4.5) below.

- (4.5.a) **Pablo vio a Lola ayer.**  
*Pablo saw Lola yesterday.*
- (4.5.b) **Pablo vio ayer a Lola.**  
*Pablo saw yesterday Lola.*
- (4.5.c) **\*La ayer vio.**
- (4.5.d) **La vio ayer.**  
*her saw(3rd sing) yesterday.*  
*(He saw her yesterday.)*

Clitics and NPs do not occur in corresponding positions in the VP. Clitics do not occur in NP positions. Examples are given in (4.6) below.

- (4.6.a) **Escribí una carta a Lola/ella/\*la.**  
*I wrote a letter to Lola/her(pronoun)/\*her(clitic).*
- (4.6.b) **Le/\*Lola/\*Ella escribí una carta.**  
*I wrote her a letter.*

The clitic promotion transformation applies only to clitics, not to NPs, as in (4.7).

- (4.7.a) **Lola quiere comerlo.**  
*Lola wants to eat it.*
- (4.7.b) **Lola lo quiere comer.**  
*Lola wants to eat it.*
- (4.7.c) **Lola quiere comer el dulce.**  
*Lola wants to eat the sweet.*
- (4.7.d) **\*Lola el dulce quiere comer.**

The sum of these arguments is very strong evidence that clitics are not NPs, and in particular are not pronominals. We could postulate a new lexical category. We take up this argument next.

#### 4.2.5. Accents and inflections

[Jaeggli 82] argues in favour of considering clitics as "*words separated from the verbs to which they are attached, as opposed to agreement inflections.*" (page 55, fn10).

To support his assertion he points out that clitics do not alter the stress pattern of their governing verbs, for example:

- (4.8.a) **cantando**  
*singing*
- (4.8.b) **cantándote**  
*singing to you*
- (4.8.c) **cantándotelo**  
*singing it to you*

Nevertheless, we can easily find true inflections that do not affect stress patterns, as in:

- (4.9.a) **sol**  
*sun*
- (4.9.b) **soles**  
*suns*

We conclude<sup>75</sup> that Jaeggli's arguments do not support the interpretation of clitics as independent words. A more likely interpretation is that clitics are inflections.

---

<sup>75</sup>Note that our schema introduced later in the Chapter treats clitics as inflections and does not change the stress pattern of the verbs involved. For example:

conjugated verb	lo pusieron (phonological stress on "e") pusieronlo pusieronlo
gerund	haciendo (phonological stress on "e") haciéndolo haciéndomelo
infinitive	lo quiere hacer (phonological stress on "e" of hacer) quiere hacerlo quiere hacértelo

For our purposes, clitics are inflections at the syntactic level. Our theory does not run counter to any phonological description of Spanish.

#### **4.2.6. Explanatory adequacy**

If clitics are not inflections then they must be an entirely new lexical category. Explanatory adequacy is improved if we do not have to add a new lexical category. It has been suggested that clitics may be pronouns (the usual textbook approach). We reject this approach for, as just illustrated above, clitics are not NPs. A simpler grammar, and thus a grammar with more explanatory adequacy, results if clitics are inflections.

#### **4.2.7. Possibilities for clitics**

Although traditionally clitics have been called weak pronouns, we have already seen that clitics do not have the same properties as pronouns. We recognise that clitics have a function similar to that of pronouns in that they are co-referent to NPs. That is, clitics, like pronouns, refer to objects or concepts and thus have the same semantic function as pronouns. Having demonstrated the fundamental differences between clitics and pronouns, we now examine clitics as elements having the properties of verb inflections.

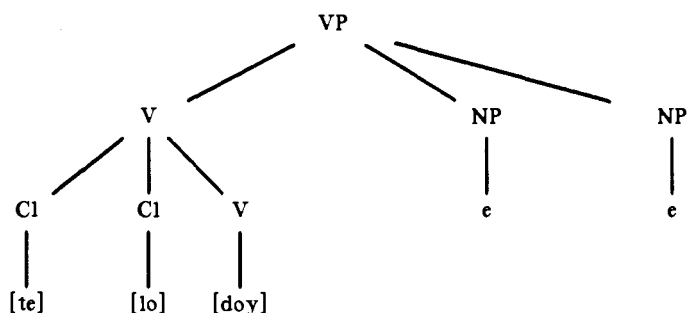
### **4.3. Base Generation of Clitics**

In this section we present briefly four examples of recent theories of the base generation of clitics in Spanish. [Rivas 77] is representative of attempts to explain clitics as an independent syntactic category generated in the base in proclitic position. [Hurtado 86] theorises that clitics are parts of syntactic chains, but still treats clitics as a syntactic category. [Safir 85] generates subject doubling clitics in VP, and finally [Roberge 85] suggests that subject doubling clitics may be generated under AGR in INFL.

### 4.3.1. Object deletion - Rivas

[Rivas 77] theorises that clitics are generated in "clitic position" and must be matched by full objects such as {NP, PP, PRO, S} as in (4.10).

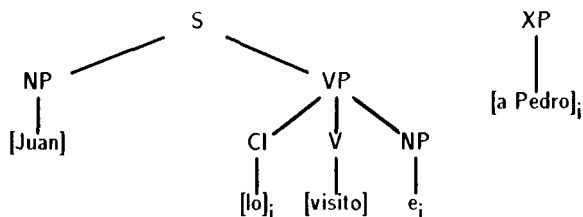
- (4.10) **Te lo doy.**  
*I give it to you.*



### 4.3.2. Clitic chains - Hurtado

[Hurtado 86] considered clitics to be chains generated in the base with the clitic in proclitic position chained to an empty NP in argument position chained in turn at the level of discourse to an extrasentential maximal projection XP. Extrasentential projections are not governed at LF and are not adjoined to some projection of S, VP, INFL, or V. XP may appear extraposed left or right of S. The XP does not receive the same  $\Theta$ -role as the clitic...e pair. This schema is illustrated in (4.11).

- (4.11) **Juan lo visitó a Pedro.** (La Plata dialect)  
*Juan visited Pedro.*

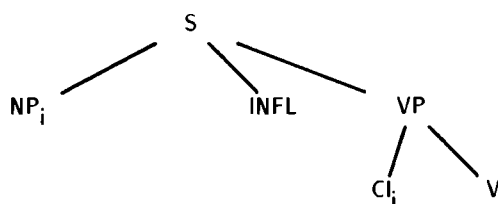




### 4.3.3. Clitic generation in VP

Another possibility derives from Safir's analysis of subject clitic doubling [Safir 85]. His analysis holds that subject clitics are generated in position in VP as a syntactic category as seen in (4.12).

(4.12)



In this analysis, the subject clitic is assigned case by INFL and the subject NP receives case from the clitic by co-indexation.

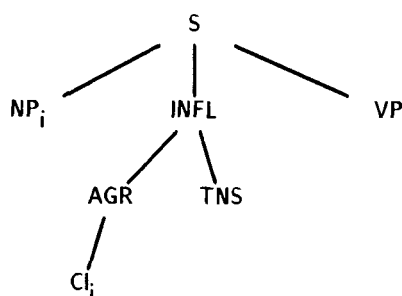
### 4.3.4. Subject clitic generation in AGR

[Roberge 85] has proposed that the subject clitic in subject doubling phenomena in Colloquial French is Base generated under AGR in INFL with the subject NP and the clitic forming a chain that receives a  $\Theta$ -role (see (4.13) below). The position occupied by the clitic is assumed<sup>76</sup> to have no status (A or  $\bar{A}$ ) with respect to the  $\Theta$ -criterion.

---

<sup>76</sup>After [Borer 81].

(4.13)



## 4.4. The Inflection Hypothesis

Now we introduce our hypothesis on the Base generation of clitics. First we introduce the notion of **enclitic1** and **enclitic2** positions. Then we introduce the notion that clitics are generated in INFL. Lastly we look at data which supports our approach.

### 4.4.1 Paradigmatic conjugation

We suggest that there are two positions for verb inflections in Spanish<sup>77</sup>, which we will call *primary inflection position* (post-position) and *secondary inflection position* (ante-position). By *primary* we mean the D-structure position for an inflection. The primary position is further subdivided into **enclitic1** and **enclitic2** positions. Thus in (4.14 a), the subject-agreement inflection occupies the primary enclitic1 position, the primary enclitic2 position is empty, and the object clitic *lo* occupies the proclitic (secondary) position. In (4.14.b) the proclitic position is empty, and the enclitic1 and enclitic2 positions are filled by the subject-agreement inflection and the object clitic respectively.

---

<sup>77</sup>In a forthcoming paper we suggest that the schema may be applied to other Romance languages with suitable parameterization.

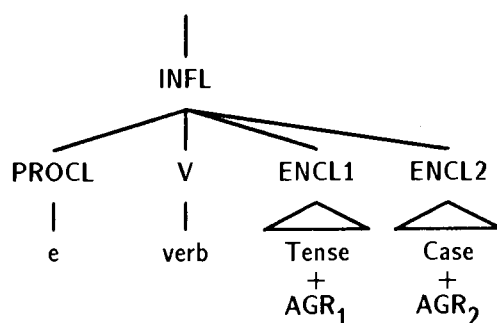
- (4.14.a) **Juan lo hizo[+TNS] para Lola.**  
*Juan made it for Lola.*
- (4.14.b) **Juan hizolo[+TNS] para Lola.** (archaic and dialectal)  
*Juan made it for Lola.*

We note that in archaic Spanish it was possible to have both subject-agreement-inflection and object-inflection (clitic) in enclitic position. This is still used dialectally (4.14.b) in the Asturian, Galacian, and Leon Spanish dialects of northwest Spain. It is this synchronic data which we attempt to explain with our theory of clitics as inflections.

#### 4.4.2. Base generation of clitics in INFL

The proposed inflection layout of **enclitic1**, **enclitic2**, and **proclitic** around a verb stem is as follows:

(4.15)

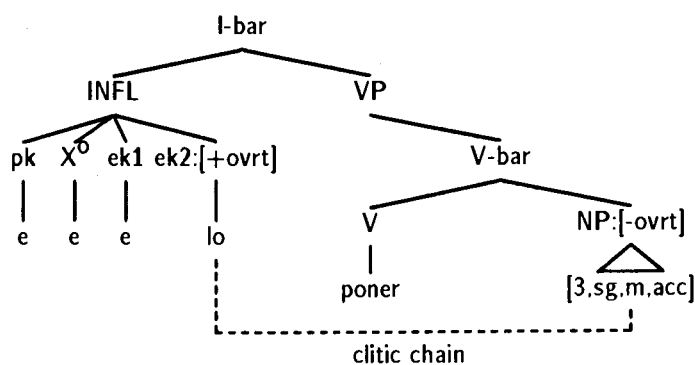


It is assumed that the Base generation of INFL provides this structure for Spanish, with PROCL empty, the verb stem position empty, the subject agreement position ENCL1 empty, and the primary clitic position ENCL2 position marked with the Case, Number, Gender, and Person information for the A-position to which the clitic is chained.

The syntactic fragment (4.16) provides an example of the hypothesised D-structure. The

verb is Base generated in VP (later to be raised to INFL). The clitic in ENCL2 is shown in its final form although this of course happens at the level of PF where the features of Case and Agreement of the clitic chain are interpreted as an inflection. We will continue to do this to ease discussion.

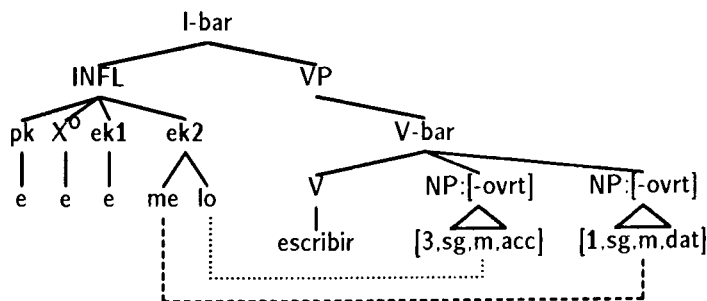
## (4.16)



In our analysis, the ENCL1 position is either filled by a subject agreement inflection (in the case of +TNS verbs) or is empty (for -TNS) verbs.

It is possible to have more than one inflection in the ENCL2 position. Consider the example in (4.17) with two clitics.

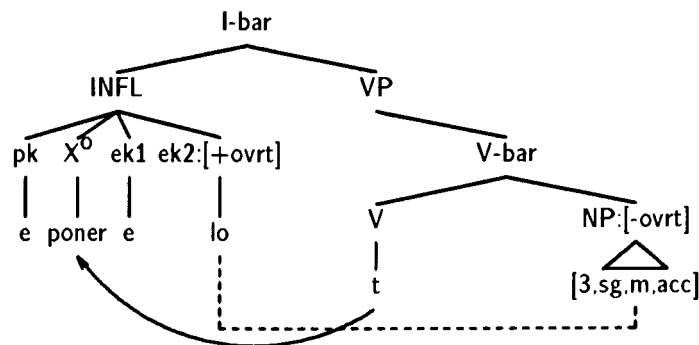
## (4.17)



### 4.4.3. Transformations involving INFL

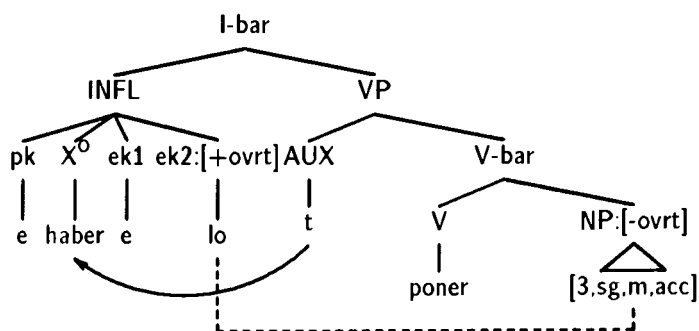
We have assumed that all Spanish clitics are generated in INFL in the Base<sup>78</sup>. There are two transformations involving the INFL structure. The first involves the raising of the verb from its D-structure position in VP to the X<sup>0</sup>-position in INFL as in (4.18).

(4.18)



Alternatively, an auxiliary may move from VP to INFL<sup>79</sup> as in (4.19). (Only one verb may move to INFL, not both.)

(4.19)

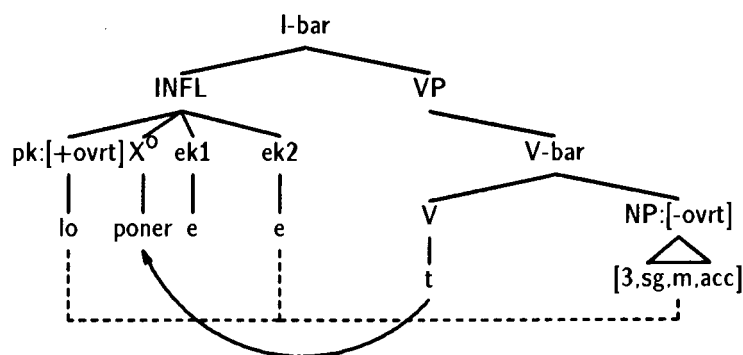


<sup>78</sup>This follows the idea espoused by Roberge [Roberge 85] for subject clitic doubling.

<sup>79</sup>The position of the Base generation of auxiliaries is the subject of some controversy. We assume here that the treatment of auxiliaries is adequate to explain the generation of clitics.

A subsequent transformation which we call **clitic hopping** (and distinct from clitic promotion) may raise the clitics from ENCL2 to PROCL under conditions to be examined in subsequent subsections, but generally based on conditions of tense. This yields the structure of (4.20).

(4.20)



We consider the generation of the clitic at PF to be purely a matter of morphological processing. Clitic hopping is not a true movement phenomenon and does not, in any case, involve a syntactic category but rather an inflection. Therefore clitic hopping does not leave a trace when it occurs. Furthermore, as inflections, clitics have no status (A or  $\bar{A}$ ) with respect to the  $\Theta$ -criterion. (A conclusion reached by Roberge. However, Roberge considered clitics to be a category and had no principled explanation of the lack of status. In our case, the lack of status is a consequence of clitics being inflections and not NPs or an independent category.)

#### 4.4.4. Imperatives

At first glance the imperative constructions in Spanish appear problematical because non-subject clitics (inflections) may occur in primary position with a tensed verb. When the source of an imperative is considered to be a complete sentence however, a possible mechanism emerges. Consider (4.21). A complete imperative sentence (4.21.a) is

generated with a subjunctive verb and the clitic in proclitic position. The middle Spanish form (4.21.b) also has the subjunctive verb and the proclitic. The modern form (4.21.d) uses an indicative form and the enclitic<sup>2</sup> position. We note that the very recent idiomatic form (4.21.c) seems to be a form intermediate to the middle Spanish form and the modern form<sup>80</sup>

We conclude that the subjunctive *used as an imperative* does not allow enclitic<sup>2</sup> to be filled. It requires hopping to proclitic as in (4.21.a,b,c).

Contrast this with non-imperative usage: dialectal as illustrated earlier, and other uses such as in traditional (4.21.e), in poetic (4.21.f), in archaic (4.21.g), and in idiomatic usage. In these cases the clitic hopping from enclitic<sup>2</sup> to proclitic is optional and **not** obligatory.

- (4.21.a) **¡Quiero que te sientes!**  
*I want you to sit down!*
- (4.21.b) **¡Que te sientes!**  
*Sit down!*
- (4.21.c) **¡Te sientes!**  
*Sit down!*
- (4.21.d) **¡Siéntate!**  
*Sit down!*
- (4.21.e) **Erase una vez . . .**  
*Once upon a time . . .*
- (4.21.f) **Después de bañarlo, depositáronlo junto a la madre.**  
(Carmen Conde, *Torre de sombra*)  
*After washing it, they put it beside the woman.*
- (4.21.g) **Llamábase don Quixote.**  
*He was called don Quixote.*

---

<sup>80</sup>This construction is, no doubt, substandard in most dialects. Currently it has become established in colloquial peninsular Spanish. The phrase *¡Se sienten, coño!* has become infamous in Spain since the abortive coup of 1982. This type of imperative was also illustrated in Roldán [Roldán 73], page 137, where she argued that the proclitic in *¡Lo hagamos!* "is not a 'real' (Roldán's quotes) imperative but a 'borrowed' one." Evidently her *ad hoc* explanation is obviated by our theory.

Data such as (4.22) are idioms. We consider such idioms to be fossilised forms. These show that the clitics have been generated archaically as enclitics.

- (4.22.a) **Erase una vez . . .**  
*Once upon a time . . .*
- (4.22.b) **¿Habrás visto?**  
*Did you see?*

Only the second person singular imperative undergoes a morphological change. Still, the clitic is able to move to proclitic position from the enclitic<sup>2</sup> position regardless of the morphological alteration, as in (4.23).

- (4.23.a) **¡Hazlo [+TNS]!**  
*Make it!*
- (4.23.b) **¡Lo hagas [+TNS]!** (dialectal)  
*Make it!*
- (4.23.c) **¡Quiero que lo hagas [+TNS]!**  
*I want you to make it!*

#### 4.4.5. Infinitives

In Spanish, the verb-complement clitic is generated in the enclitic<sup>2</sup> position of an infinitive. (The enclitic<sup>1</sup> position is empty, and not filled with a subject-agreement clitic.)

- (4.24.a) **Puedo hacerlo.**  
*I can make it.*
- (4.24.b) **Lo puedo hacer.**  
*I can make it.*
- (4.24.c) **Isabel no quiere dármela.**  
*Isabel does not want to give me it.*
- (4.24.d) **Isabel no me la quiere dar.**  
*Isabel does not want to give me it.*

We associate this with the [+TNS] feature of clitics and note that when a clitic is promoted from enclitic of an infinitive (4.25.c) to a tensed verb, it cyclically takes the enclitic<sup>2</sup> position (4.25 d), then the proclitic position as expected (4.25.e).



- (4.25.a) **Es fácil hacérselo.**  
*It is easy to make it for her.*
- (4.25.b) **\*Es fácil se lo hacer.**
- (4.25.c) **Quiero [+TNS] hacerlo [-TNS].**  
*I want to make it.*
- (4.25.d) **Quiérollo hacer.** (dialectal and literary)  
*I want to make it.*
- (4.25.e) **Lo quiero [+TNS] hacer[-TNS].**  
*I want to make it.*

#### 4.4.6. Gerunds

We view gerunds as untensed verbs. Thus we would expect to find clitics in primary enclitic2 position (4.26.a) and not in secondary (proclitic) position (4.26.b).

- (4.26.a) **Haciéndoselo la harás feliz.**  
*Making it for her you will make her happy.*
- (4.26.b) **\*Se lo haciendo la harás feliz.**

#### 4.4.7. Past participles

Likewise clitics always need a [+TNS] auxiliary verb when they occur in the environment of past participles. In my analysis, clitic generation takes place in INFL and the auxiliary verb raises to the X<sup>0</sup> position in INFL.

- (4.27.a) **Lo ha roto.**  
*He broke it.*
- (4.27.b) **Halo roto.** (dialectal)  
*He broke it.*
- (4.27.c) **\*Ha rotolo.**

#### 4.4.8. Clitic promotion

Clitic promotion can be accounted for (syntactically) as follows. In our analysis, the clitic moves from the enclitic2 position of the infinitive verb for which it is a semantic argument to a clitic position of a higher verb. The movement is cyclical until the clitic is moved to the proclitic position of a tensed verb.

- (4.28.a) **Quería seguir gritándomelo**<sup>81</sup>.  
*He wanted to follow shouting it to me.*
- (4.28.b) **Quería seguirmelo gritando.**  
*He wanted to follow shouting it to me.*
- (4.28.c) **Queríamelo seguir gritando.**  
*He wanted to follow shouting it to me.*
- (4.28.d) **Me lo quería seguir gritando.**  
*He wanted to follow shouting it to me.*

Our treatment of clitic promotion follows from the rules of clitic placement with tensed and untensed verbs explained above and from an adaptation of Hurtado's **clitic chains** [Hurtado 86] to our theory of clitics as inflections. Because clitics are in our view not a syntactic category we consider that clitic promotion leaves no trace of the clitic when clitic promotion occurs.

#### 4.5. Some Observations

One obvious question is: why are clitics written as distinct lexical items when they precede the verb? We concur with [Strozer 76] that this is an orthographic convention, and perhaps a historical accident. The fact is, clitics are never separated phonetically from the (syntactically) associated verb. This is surely associated with the fact that clitics are never stressed. We take the lack of stress and the phonetic slurring as further evidence that clitics are inflections on the verb.

None of the arguments used above in support of clitics as verb inflections are unique to Spanish, nor are the mechanisms hypothesised for Spanish untenable for other Romance languages. There are no reasons we are aware of to treat clitics as a lexical category.

---

<sup>81</sup>(4.28.a.b.d) are from Perlmutter 1971 [Perlmutter 71].

## Chapter 5

### Generating D-structure for Spanish

The generation of D-structure from the Conceptual Graph semantic representation of a proposition involves the Lexicon and the Categorical Component of the grammar. *Base generation* is the generation of D-structure by these two components. In traditional Government and Binding treatment, free generation of sentences by the Base component produces all possible sentences. Other components and principles of the grammar then serve to constrain the generated sentences, that is, rule out some ungrammatical ones. In this computational model of NL generation, a semantic representation of the single sentence to be generated is assumed. The Environment of Syntax (EOS) contains the semantic representation (and other components discussed in Chapter 3). Procedures that provide the interface between the EOS and the Base grammar, called **consultcontext** procedures, were introduced there. This Chapter gives a more detailed view of the D-structure generation of Spanish sentences for the following specific constructions:

- Impersonal-se constructions.
- Benefactive dative constructions.
- True passive constructions.

All of these are examples of special requirements in the generation of D-structure. The impersonal-se construction requires that the grammar check for an unspecified agent when the EOS is consulted. The benefactive dative requires pre-lexical processing to introduce a new external argument of the verb and to assign the agent to an internal argument position. The true passives require pre-lexical processing and a specific lexical entry for

verbs that passivise. Generation of D-structures for these constructions is done through pre-lexical processing in the EOS consultations. Those processes are explained below.

We start with a description of the Lexicon and the Categorical Component of the grammar, which together provide Base generation of D-structures. The role of Case Theory and Theta Theory in this component is also discussed.

## 5.1. X-Bar Theory

We saw in Chapter 3 the basis in X-bar theory of the fundamental phrase structure rule of the grammar. Rules of the general form (5.1) were given. Now we want to look at the nature of the informant lists,  $\text{Inf}_x$ .

(5.1)

$\text{xp}(\text{Cat}, \text{PT}, \text{Graph}, \text{Inf}) \rightarrow \text{spec}(\text{Cat}, \text{PT0}, \text{Graph}, \text{Inf0}), \text{xbar}(\text{Cat}, \text{PT1}, \text{Graph}, \text{Inf1}).$

$\text{xbar}(\text{Cat}, \text{PT}, \text{Graph}, \text{Inf}) \rightarrow \text{x}(\text{Cat}, \text{PT0}, \text{Graph}, \text{Inf0}), \text{compl}(\text{Cat}, \text{PT1}, \text{Graph}, \text{Inf1}).$

where

**Cat** stands for the syntactic category.

**PT** is an internal parse-tree structure being built up out of **PT1** and **PT0** from the subcategories.

**Graph** is the conceptual graph from which the sentence is being generated.

**Inf** is the informant list of features; some being passed on to subsequent rules, some to be unified with unbound informants.

**Inf0** and **Inf1** are informant lists of the subprocedures.

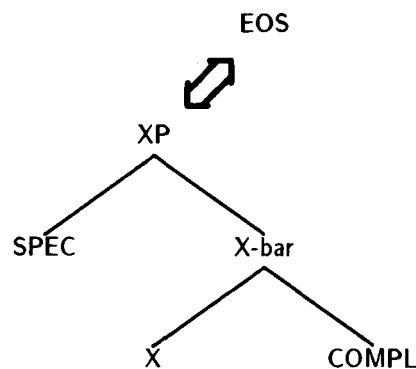
Elements of all the informant lists may be unified.

Rules were augmented (as in 5.2) with Prolog calls to the **consultcontext** procedures for each **XP** as illustrated graphically in (5.3).

(5.2)

$\text{xp}(\text{Cat}, \dots, \text{Graph}, \dots) \rightarrow$   
 $\{\text{consultcontext}(\text{Cat}, \text{Graph}, \dots)\},$   
 $\text{spec}(\text{Cat}, \dots),$   
 $\text{xbar}(\text{Cat}, \dots).$

(5.3)



We will now examine such rules in more detail. In order to do this, the subcategorisation mechanism of the grammar must be understood. In GB theory, the complement category (**COMPL**) represents the subcategorisation of the **XP** being generated. Because the **COMPL** is itself an **XP**, and because the nature of the complement-XP is determined by the lexicon (in the case of lexical categories), we turn first to the Lexicon.

Lexical entries contain subcategorisation frame information. For verbs and prepositions, this consists of the subcategorisation category together with  $\theta$ -role and Case features for the external and internal arguments. In the example entry (5.4), the internal argument (the complement category) of the verb is **np** with the feature list **[theme,acc,Def2|R2]**, while the external argument has the feature list **[agt,nom,Def1|R1]**.

(5.4) **verb(need1,necesitar,np,[[agt,nom,Def1|R1],[theme,acc,Def2|R2]])**.

When the VP rule (5.5) is called, the `consultcontext` subprocedure unifies **SubCat** with **np**, **SubjInf** with **[agt,nom,Def1|R1]**, and **ComplInf** with **[theme,acc,Def2|R2]** through

its interaction with the EOS<sup>82</sup>.

(5.5)

**xp(vp,PT,G,[TNS,actv,-bnft,SubjInf,ComplInf]) -->**  
**{consultcontext(vp,W,G,SubCat,Voice,-bnft,[SubjInf,ComplInf]),**  
**spec(vp,PT1,G,[Voice]),**  
**xbar(vp,PT2,G,W,SubCat,ComplInf),**  
**{PT=..[vp,[TNS],PT1,PT2]}.**

Similarly, through the consultation with the EOS, the subcategory of the COMPL is found for all XPs. The subcategorisation frame for lexical entries is a language specific feature of the lexicon. It allows the grammars of different languages to express the same meaning through different syntactic structures.

The rule (5.5) is a complete rule of the Base grammar. (It differs from the incomplete rule (5.2) in having an extra call (**{PT=..[vp, . . .]}**) to a Prolog procedure that constructs the abstract parse tree representation of the D-structure and in having a complete set of informants.)

We now have the general form of the complete rules for generating the basic XP syntactic structure in (5.6). The informant lists (**Inf<sub>x</sub>**) are shown in a very general form. They are discussed below for specific categories.

---

<sup>82</sup>The informant **-bnft**, the benefit feature, will be discussed in the section on generating benefactive dative constructions.

(5.6)

```

xp(Cat,PT,Graph,Inf)-->
    {consultcontext(Cat,Word,Graph,SubCat,Inf0)},
    spec(Cat,PT0,Graph,Inf1),
    xbar(Cat,PT1,Graph,Inf2),
    {PT=..[Cat,Inf3]}.
xbar(Cat,PT,Graph,Inf)-->
    x(Cat,PT0,Graph,Inf0),
    compl(SubCat,PT1,Graph,Inf1),
    {PT=..[Cat,Inf2]}.

```

The rules for specifiers and complements need some explaining. (5.6) has category-specific calls to procedures to generate specifiers and complements. The procedure to generate a specifier (**spec**) may call a procedure to generate an XP (such as the NP specifier of IP) or a procedure to generate a determiner (for the specifier of an NP), or a procedure to generate an empty XP position (such as the specifier of CP). Similarly, the procedure to generate the complement (**compl**) will call a procedure to generate an XP of the **SubCat** category

(5.7)

```

spec(ip,PT,Graph,[actv,-imp,-bnft,SubjInf]) --> xp(np,PT,Graph,SubjInf).

spec(np,PT,Graph,[Def,Agr,+ovrt]) --> det(PT,[Def,Agr,+ovrt]).

spec(cp,PT,Graph) --> {PT=xp([],e)}, [e].

compl(SubCat,PT,Graph,Inf) --> xp(SubCat,PT,Graph,Inf).

```

The informant lists that carry syntactic features and other information between procedure calls are important. They will be treated below for individual categorial types. How Case theory and Theta theory are obeyed through the use of these informants is also discussed below.

## 5.2. Case and Theta Theory in D-structure

Case Theory in GB states simply that: *Every phonetically realised NP must have Case assigned*. Case may be assigned in either of two ways. Inherent Case marking applies in the Base component. Structural Case marking applies at S-structure. In addition, Case filters operate to rule ungrammatical some constructions<sup>83</sup>. As we have seen, Case is assigned to A-positions in D-structure by projection from the Lexicon as the EOS is consulted from VP.

The Theta Criterion states that: *Each argument bears one and only one  $\Theta$ -role*.  $\Theta$ -roles are also assigned to A-positions in D-structure by projection from the Lexicon as the EOS is consulted from VP. The Projection Principle may rule out some sentences as ungrammatical, based on violations of the Theta Criterion<sup>84</sup>. These principles of GB theory are the reason the informant lists seen in the previous sections carry Case and  $\Theta$ -role features

## 5.3. Getting Features from the EOS

The interaction of Lexicon and Conceptual Graphs in the EOS to get or verify  $\Theta$ -roles, Case features, and subcategorisation categories was discussed in Chapter 3. That discussion will not be repeated here. However, the informant lists ( $\text{Inf}_x$ ) carry some other semantic and syntactic features particular to the category of the XP being generated. The roles of those informant lists are explained in the following subsections.

---

<sup>83</sup>The filters (The NP-trace filter, the PRO filter, the Case filter, and the Case Conflict Filter) will concern us when we consider the transformational component in Chapter 7.

<sup>84</sup>In Chapter 7 we will be concerned whether the Theta Criterion is obeyed after transformations.



The effect of the grammar is to achieve projection of  $\Theta$ -role, Case and other features from VP to the arguments of the verb. This is actually achieved in a top-down generation through the non-determinism of Prolog. Features are carried down the derivational context as unbound variables until they can be unified with a particular value. If, later in the derivational context, the bound value is found to be incompatible with further generation, backtracking of the Prolog interpreter will undo the binding and allow another binding to take place.

### 5.3.1. Informants at the CP level

The informant list for CP contains [**Level**,**Voice**,**Imp**,**Type**,**TNS**,**TR**] as a prototypical list that may be set at the clausal level of generation. **Level** indicates whether the CP is a matrix clause or an embedded clause. **Voice** say whether the CP is active or passive. **Imp** is a feature indicating whether the CP is or is not an impersonal construction and takes one of the values  $\{\pm\text{imp}\}$  **Type** indicates whether the CP is declarative, imperative, or interrogative. **TNS** is a feature indicating whether or not the CP is tensed or not and thus takes on one of the values  $\{(\pm)\text{tns}\}$ . Finally, **TR** has been included as the  $\Theta$ -role of the CP when appropriate

Of these informants, **Level** and **Type** are bound at the time of the call to generate CP. **Voice** is unbound to allow free generation of passives. **Imp** is projected from the VP, and **TNS** is currently set by default to  $+\text{tns}$  but could be projected from IP.

(5.8)

```

xp(cp,PT,G,[Level,Voice,Imp,Type,TNS,TR]) -->
    {consultcontext(cp,W,G,SubCat,[Level,Voice,Type,TR])},
    spec(cp,PT1,G),
    xbar(cp,PT2,G,W,SubCat,[Level,TNS,Voice,Imp]),
    {PT=..[cp,[Level,Voice,Imp,Type],PT1,PT2]}.

xbar(cp,PT,G,W,SubCat,[Level,TNS,Voice,Imp]) -->
    x(cp,PT1,G,W,[Level,TNS]),
    compl(SubCat,PT2,G,[TNS,Voice,Imp]),
    {PT=..[cbar,PT1,PT2]}.

```

### 5.3.2. Informants at the IP level

At the call to generate IP, the informant list is [TNS,Voice,Imp,Bnft]. Imp is unified at the SPEC of IP level with [+imp] if an impersonal subject is to be generated, and [-imp] otherwise. The Bnft informant is unified at the VP level and projected to IP in order to generate the impersonal subject necessary for the benefactive dative construction. An additional informant introduced by the consultation with the EOS for IP is Tense, which is unified with the actual tense (for example, past) and used in the generation of INFL. The subject informant list, SubjInf, is projected to the level of the spec(ip, . . .) call from VP, and both the subject informant list and the complement informant list(s) are projected to the xbar(ip, . . .) call.

(5.9)

```

xp(ip,PT,G,[TNS,Voice,Imp,Bnft]) --->
    {consultcontext(ip,W,G,SubCat,[TNS])},
    spec(ip,PT1,G,[Voice,Imp,Bnft,SubjInf]),
    xbar(ip,PT2,G,SubCat,[TNS,Voice,Imp,Bnft,SubjInf]),
    {PT=..[ip,[TNS,Imp,Bnft],PT1,PT2]}.

xbar(ip,PT,G,SubCat,[TNS,Voice,Imp,-bnft,SubjInf]) -->
    x(ip,PT1,G,[TNS,Voice,Imp,Bnft,ObjInf]),
    compl(SubCat,PT2,G,[TNS,Voice,Bnft,SubjInf,ObjInf]),
    {PT=..[ibar,PT1,PT2]}.

```

### 5.3.3. Informants at the VP level

The informants at the VP level are **[Voice,Bnft,TNS,SubjInf,ComplInf]**. Of these, **Voice** has been unified at the clausal (CP) level and has been carried down in informant lists to the VP level to determine whether pre-lexical passive processing is necessary. **Bnft** has been unified at the IP level and has been carried down to the VP to determine whether pre-lexical processing for the benefactive dative construction is necessary.

**SubjInf** is the list of informants of the subject A-position. The TR ( $\theta$ -role) and Case of this list are projected from the VP. (That is, they are unified at the VP level by the **consultcontext(vp, . . .)** call.) Similarly, **ComplInf** is a list of informants for the object A-position. TR and Case are projected from VP here also. If the verb found in the Lexicon by **consultcontext(vp, . . .)** is a di-transitive verb, then there are two complement informant lists called **DObjInf** and **IObjInf**.

(5.10)

```

xp(vp,PT,G,[TNS,actv,-bnft,SubjInf,ComplInf]) -->
    {consultcontext(vp,W,G,SubCat,Voice,-bnft,
      [SubjInf,ComplInf])},
    spec(vp,PT1,G,[Voice]),
    xbar(vp,PT2,G,W,SubCat,ComplInf),
    {PT=..[vp,[TNS],PT1,PT2]}.

xbar(vp,PT,G,W,SubCat,ComplInf) -->
    x0(vp,PT1,G,W),
    compl(SubCat,PT2,G,ComplInf),
    {PT=..[vbar,PT1,PT2]}.

```

### 5.3.4. Informants at the NP level

The informant list for NP is [TR,Case,Agr,Def,Ovrt,Ani]. As seen earlier, TR and Case are projected to the A-position from the VP. Agr, Def, and Ani are unified by the `consultcontext(np, . . .)` consultation with the EOS. Ovrt is bound to {±ovrt} by the grammar according to rules of dialect, emphasis, Case, and pronominalisation as outlined in Chapter 3.

As part of a `SubjInf` or `ComplInf` list, the Case, Ovrt, Agr, and Def features are allowed to percolate up or filter down to INFL (through the unification process of Prolog). Case, Ovrt, and Agr play a role in determining the overtiness and/or the form of the clitic chained to the A-position of the NP

(5.11)

```

xp(np,PT,G,[TR,Case,Def,Agr,+ovrt,Animate]) -->
    {consultcontext(np,Word,G,SubCat,[Def,Agr,TR,Animate])},
    spec(np,PT1,[Def,Agr,+ovrt]),
    xbar(np,PT2,G,Word,SubCat,+ovrt),
    {PT=..[np,[TR,Case,Def,Agr],PT1,PT2]}.

```

```

xp(np,PT,G,[TR,Case,Def,Agr,-ovrt,Animate]) -->
    {consultcontext(np,Word,G,SubCat,[Def,Agr,TR,Animate])},
    [-ovrt-np],
    {PT=..[np,[TR,Case,Def,Agr],nil]}.

```

```

xbar(np,PT,G,Word,SubCat,+ovrt) -->
    x(np,PT1,Word,+ovrt),
    compl(SubCat,PT2,G,Informants),
    {PT=..[nbar,PT1,PT2]}.

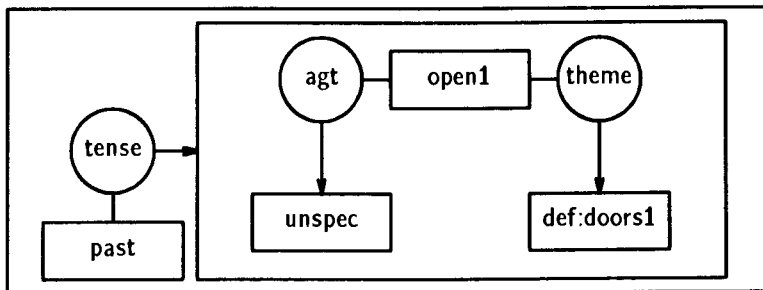
```

The second rule for `xp(np, . . .)` in (5.11) generates NPs with the [-ovrt] feature. That is phonetically empty NPs chained to a clitic inflection.

## 5.4. Generating Impersonal-se D-structure

These constructions involve an unspecified subject at the level of the semantic representation as we saw in (2.16):

- (2.16.a) **Se abrieron las puertas.**  
 (2.16.b) **Se abrió las puertas.** (dialectal usage in Spain)  
*The doors were opened (by somebody).*



In Spanish this meaning is usually realised as a clitic chain with a **pro**<sup>85</sup> in the subject A-position and the impersonal subject clitic **se**.

- (5.12)     **pro**     . . .     **se**  
              [-OVRT]                [+OVRT]

The rules for the SPEC of IP (the subject A-position) must therefore check whether a subject is unspecified or not. First a call to generate a [-ovrt]NP with the appropriate agreement features is made which verifies the presence of the impersonal subject (that is the unspecified agent) in the EOS. Then the feature IMP may be unified with [+imp]. The agreement features are assumed to be **agr(3,Nbr,n)**. Note that the number feature,

<sup>85</sup>**pro** is the phonologically empty pronominal element in Linguistics.

**Nbr** is unbound as yet<sup>86</sup>, and that the gender feature is unified with **n** (specifying no gender assigned).

(5.13)  
**spec(ip,PT,Graph,[Voice,+imp,-bnft,[agt,nom,null,agr(3,Nbr,n),-ovrt,Ani]])**  
 -->  
**xp(np,PT,Graph,[agt,nom,null,agr(3,Nbr,n),-ovrt,Ani]),!**

This call to generate an NP with very specific features is passed to a rule for impersonal subject NPs so that the EOS consultation is a verification of the unspecified subject and the agreement features:

(5.14)  
**xp(np,PT,Graph,[agt,nom,null,agr(3,Nbr,n),-ovrt,Ani]) -->**  
**{consultcontext(np,un,Graph,[],[null,agr(3,Nbr,n),agt,Ani]}**  
**!,[pro],**  
**{PT=..[np,[agt,nom,null,agr(3,Nbr,n)],pro]}.**

Corresponding lexical entries are the noun entry (5.15.a) called by **consultcontext** and the clitic entry (5.15.b) generated in the ENCL2 position of INFL.

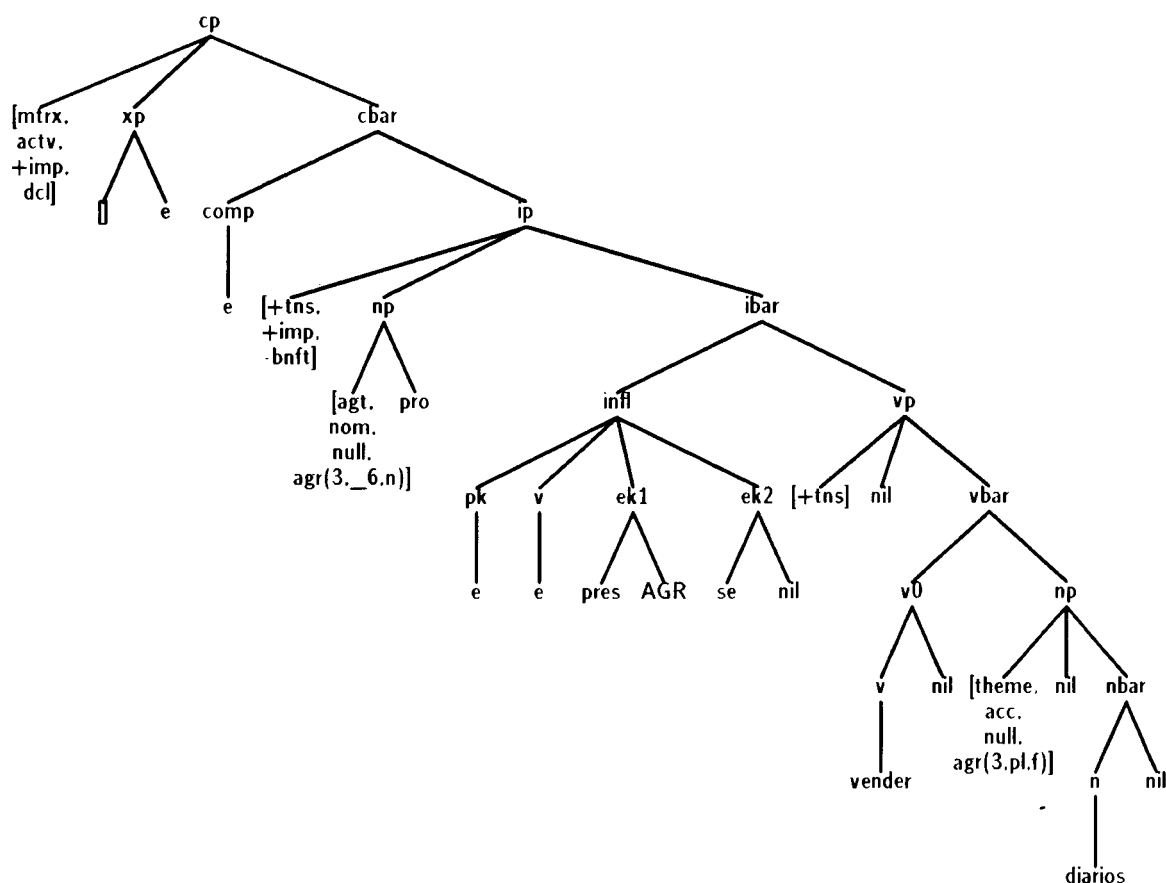
(5.15.a) **noun(unspec,un,[],agr(3,Nbr,n),entity).**  
 (5.15.b) **clitic(se,Case,+ovrt,agr(3,Nbr,n)) --> [se].**

(5.16) shows the actual internal D-structure representation generated for such a sentence.

---

<sup>86</sup>It is beyond the scope of this dissertation to do the late assignment of number to the verb and the unspecified subject.

## (5.16) Se venden diarios. (D-structure)



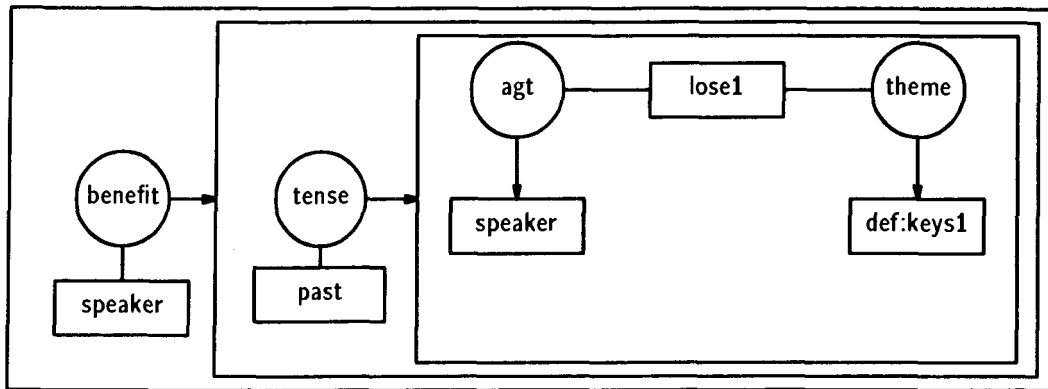
## 5.5. Generating Benefactive Dative D-structure

Recall that the benefactive dative constructions were typified by a **benefit** relation<sup>87</sup> involving the **agent** at the level of semantic representation.

In the example (2.31), repeated here, and in all cases of the benefactive dative, the sentence has a verb with one more argument position than the semantic representation

<sup>87</sup>Recall that the **tense** relation should perhaps be given wider scope over the **benefit** relation according to ter Muelen's view of the semantics of tense and aspect mentioned in Chapter 2.

- (2.31) **Se me perdieron las llaves.**  
*The keys got lost on me.*



indicates. This may be typified as a mono-transitive verb in the lexicon becoming a di-transitive verb in the D-structure.

There are two ways to handle the generation of such sentences. One is to treat the benefactive dative as a lexically governed construction (as the passive construction will later be treated). The other is to posit a pre-lexical processing of the proposition to add a new argument. The latter approach is taken here because of the extremely productive nature of the benefactive dative construction in Spanish. To do otherwise would mean that lexical entries for benefactive dative verbs would have to be made corresponding to almost all action verb entries, essentially doubling the number of lexical entries for verbs, with no apparent gain in explanatory power. On the other hand, to postulate and demonstrate a single pre-lexical process that generates the benefactive dative, whenever the semantic representation allows it, has great explanatory power.

The unspecified subject of this construction is added to the argument structure present in the semantic representation through a special call to the EOS from the generation of the VP level of syntax: In (5.17) the BNFT feature is unified with (+bnft) and a new



external A position with  $\Theta$ -role **experiencer** is created (and projected from VP to IP) while the **processgraph** call has only to verify the two  $\Theta$ -roles **agent** and **theme** in the semantic representation.

(5.17)

```
consultcontext(vp,Word,Graph,[np,SubCat],Voice,+bnft,
  [[experiencer,nom,null,agr(3,Nbr,n),-ovrt, Animate0],
   [theme,acc,Def2,Agr2,Overt2,Animate2],
   [agt,dat,Def1,Agr1,-ovrt, Animate1]])
:- processgraph(Graph,Token,[agt,theme]),
  verb(Token,Word,SubCat,[[agt|Rest1],[theme|Rest2]]),
  consultstack,checkexternal,constraints.
```

Note that the agreement features of the **experiencer** are **agr(3,Nbr,n)**, as expected for an impersonal subject construction.

The generation of the SPEC of IP projected from the VP introduces the impersonal subject, not by consulting the EOS (which would fail of course, there being no unspecified subject in the semantic representation), but directly.

(5.18)

```
spec(ip,PT,G,[Voice,+imp,+bnft,
  [experiencer,nom,null,agr(3,Nbr,n),-ovrt,animate]])
--> {consultcontext(subject,_,G,+bnft,[benefit])},
  !, [pro],
  {PT=..[np,[experiencer,nom,null,agr(3,Nbr,n)],pro]}.
```

The call to the EOS here involves verification that the **benefit** relation on the **agent** of the proposition exists:

```
consultcontext(subject,_,Graph,+bnft,[benefit]) :-
  processgraph(Graph,agent,benefit), !.
```

The agreement features of the **experiencer** (impersonal subject) are projected to the ENCL2 position resulting in the clitic chain.

## (5.19)

<b>pro</b>	. . .	<b>se</b>
[-OVRT]		[+OVRT]

Returning to the other  $\Theta$ -roles of the verb, we see that the **theme** is assigned accusative (**acc**) case, and the **agent** is assigned dative (**dat**) case but marked [-ovrt]<sup>88</sup>. Thus the external argument (the **agent**) in the lexicon becomes an internal argument in the syntax.

The parallels between Perlmutter's analysis [Perlmutter 87] of the Causative constructions in French, which suggests a pre-lexical<sup>89</sup> level of processing for several constructions, and that suggested above for the benefactive dative are notable. In his Grammatical Relations paradigm, Perlmutter determines the grammatical relation of the argument of a predication with respect to a set of principles and constraints<sup>90</sup>. A subject may become an indirect object under a set of 'promotion' and 'demotion' processes which are subject to the principles constraints.

---

<sup>88</sup>Some native speakers of Spanish insist that the A-position of the benefactive dative clitic chain may not be marked [+ovrt]. I have taken the point of view that this is the most common usage. It would be easy to let the grammar generate a [+ovrt] NP for this dative through backtracking.

**Se le perdieron las llaves.**

*His keys got lost on him.*

**Se le perdieron las llaves a Juan.**

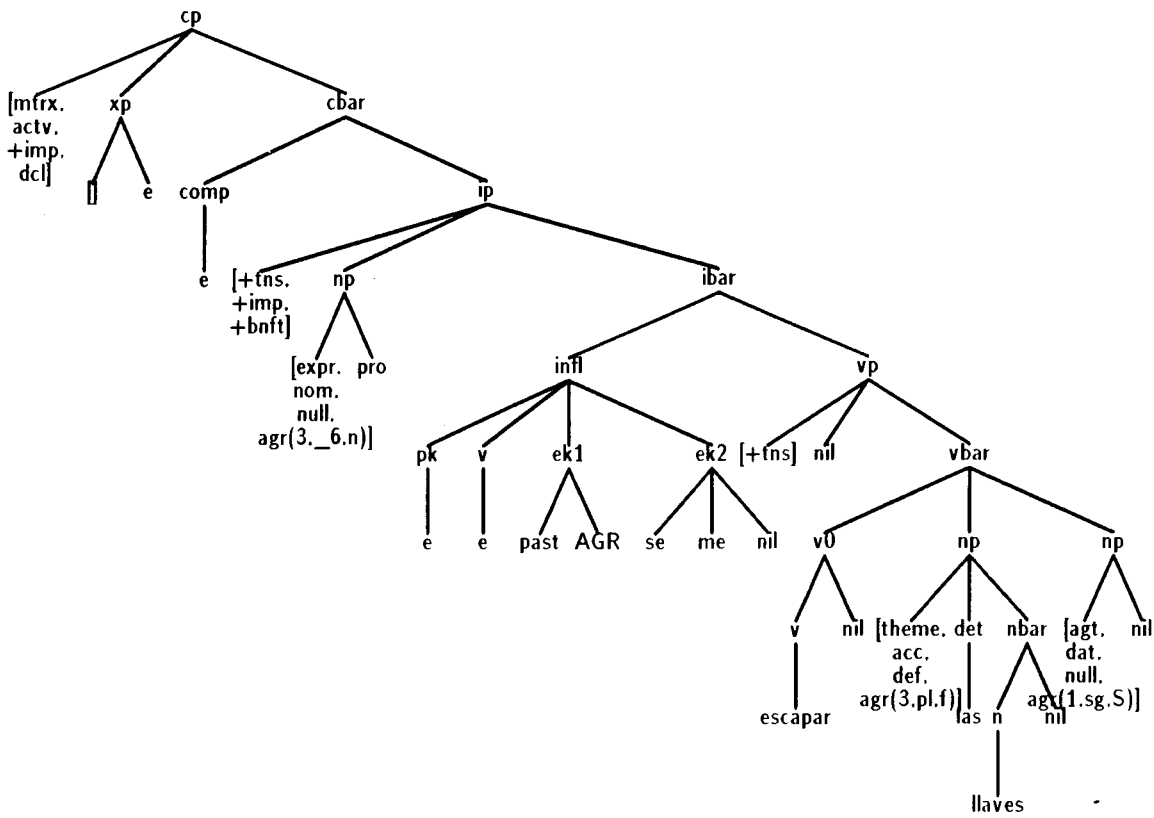
*Juan's keys got lost on him.*

In this case, the **agent** should be given the unbound value [OVRT]. To the best of my knowledge, this insistence on no overt NP by some speakers has not been discussed previously in the corpus of literature on Spanish clitics.

<sup>89</sup>Perlmutter's term.

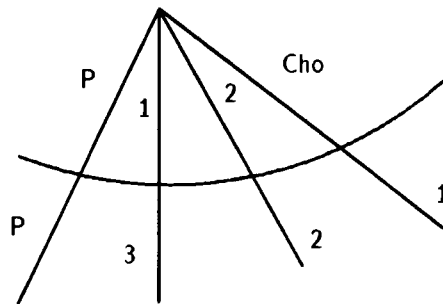
<sup>90</sup>The principles and constraints include: the stratal uniqueness law, the successor demotion ban, the chomeur advancement ban, and the oblique law.

(5.20) Se me perdieron las llaves. (D-structure)



The diagram (5.21) illustrates a similar pre-lexical process in Perlmutter's formalism.

(5.21)



I note in passing that all of Perlmutter's processes can be considered as a simultaneous set of processes. For example:

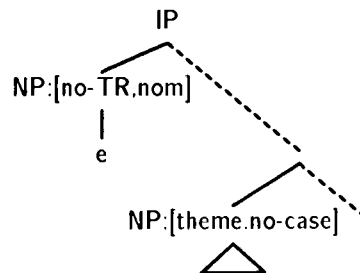
(5.22)		
Subject	<i>becomes</i>	Indirect Object
Direct Object	<i>becomes</i>	Direct Object
Chomeur	<i>becomes</i>	Subject

This paradigm is therefore suitable for processing by CDG rules which express exactly this kind of simultaneous processing.

## 5.6. Generating True Passive D-structure

Passives are assumed to be generated from a D-structure with an empty XP position in the SPEC of IP and an NP without Case in an internal A-position as in (5.23). (In Lectures in Government and Binding, it was proposed [Chomsky 82a] that passive verbs do not assign Case.)

(5.23)

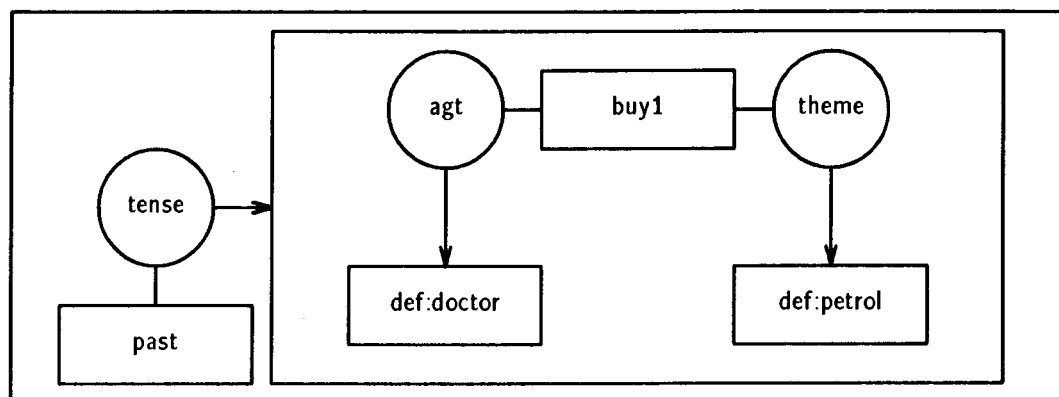


Move- $\alpha$  subsequently moves the NP without Case into the SPEC of IP position. This transformation is discussed at length in Chapter 7. Our interest here lies in the generation of the D-structure for the active and true passive versions of a sentence.

Consider a semantic structure such as that given in (1.6) repeated here as (5.24).

The two sentences have different D-structures. That of the active (5.24.a) corresponds directly to the argument structure of the semantic representation. That of the passive

- (5.24.a) **El médico compró la bencina.**  
*The doctor bought the petrol.*
- (5.24.b) **La bencina fue comprada por el médico.**  
*The petrol was bought by the doctor.*



(5.24.b) does not correspond to the argument structure of the semantic representation, but rather to an argument structure imposed by the lexical entry for a verb that may be passivised

Verbs which may passivise have two entries in the Lexicon. In the the case of the verb **comprar**, these are the entries of (5.25). The second of these is the entry which is used for passive D-structure.

(5.25)  
**verb(buy1, comprar, np, [[agt,nom,Def1|Rest1],  
 [theme,acc,Def2|Rest2]]).**

**verb(buy1, comprar, [np,pp(por,np)], [[no-tr,nom,Def1|Rest1],  
 [theme,nocase,Def2|Rest2],  
 [agt,obl,Def3|Rest3]]).**

A corresponding consultation with the EOS is required to assign Case and the  $\Theta$ -roles to the D-structure in a manner compatible with the relations of the semantic representation. This again is a type of pre-lexical processing that adds a new argument with no  $\Theta$ -role

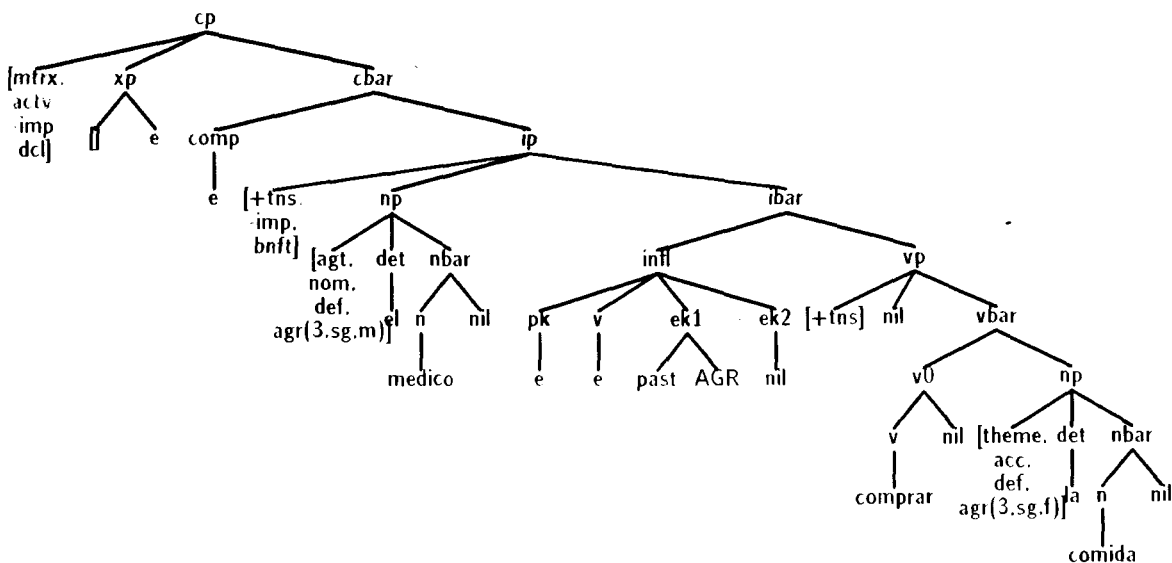
(marked **no-tr**) to the syntactic structure but in this case the pre-lexical processing is imposed by the lexical entry of the verb.

(5.26)

```
consultcontext(vp,Word,Graph,SubCat,pass,-bnft,
               [[no-tr|Rest1],[ThemeTR|Rest2],[AgtTR|Rest3]]) :-
  :- processgraph(Graph,Token,[AgtTR,ThemeTR]),
     verb(Token,Word,SubCat,[[SubjTR|Rest1],
                              [ThemeTR|Rest2],
                              [AgtTR|Rest3]]).
```

The resulting D-structures are found in (5.27) and (5.28) for the active and passive versions of the sentence respectively.

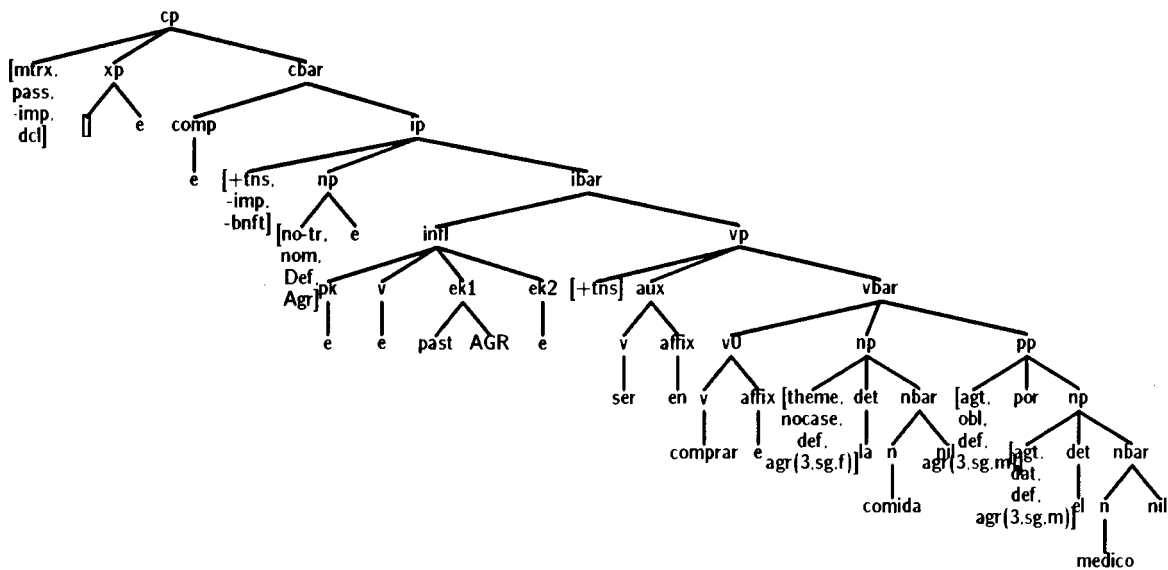
(5.27) El médico compró la comida. (D-structure)



## 5.7. Non-deterministic Generation

The generation grammar will generate all possible sentences through the non-deterministic backtracking capability of Prolog. For example, when free clitic substitution for NPs is permitted, D-structures for the following sentences (corresponding to the semantic structure (16)) were generated.

## (5.28) La comida fue comprada por el médico. (D-structure)



- (5.29.a) El médico la compró.  
*The doctor bought it.*
- (5.29.b) El médico compró la comida.  
*The doctor bought the food.*
- (5.29.c) La compró.  
*He bought it.*
- (5.29.d) Compró la comida.  
*He bought the food.*
- (5.29.e) La comida fue comprada por el médico.  
*The food was bought by the doctor.*
- (5.29.f) La comida fue comprada.  
*The food was bought.*

## 5.8. Comments

The D-structure generation of the grammar make use of DCG type rules which are later compiled into CDG rules. The EOS has been demonstrated to be effective in Base generating clitic chains according to the theory of clitics as inflections. The use of informant lists as a means of implementing projection of features from VP has been successful. The unification facility of Prolog makes this possible in a top-down generation system based on X-Bar Theory.

The combination of consultation with the EOS for each XP generated and the ability to introduce pre-lexical processing in the consultation has resulted in an extremely powerful generation paradigm in itself. The generation of the D-structures for impersonal-se, beneficial datives, and true passives has been demonstrated here. In chapter 7 the additional power of transformations is introduced through Constrained Discontinuous Grammar rules.



## Chapter 6

### A New Formulation for Discontinuous Grammars

Discontinuous Grammars (DGs) [Dahl and Abramson 84] were originally conceived as a generalisation of the Extraposition Grammars (XGs) of [F. Pereira 81]. This Chapter investigates the adequacy of DGs for NL generation. The conclusion reached is that there are insurmountable problems with DGs as originally conceived in this application. An investigation of the linguistic requirements leads to a new interpretation of Discontinuous Grammars that has the following properties

- The derivational context is a true tree and not a graph as in the original formulation.
- Because of the tree form, *Flattening* of the derivational context does not occur
- Both analysis and generation of NL are possible using the new formulation, unlike the original formulation, where only analysis is possible.
- The procedural semantics of the formalism is *goal substitution* instead of the goal re-ordering of the original formulation.
- The declarative semantics is of parallel rewriting rules. This, it turns out, has interesting applications in parallel processing programming problems [Dahl and Brown 86], the pursuit of which is beyond the scope of this dissertation.

#### 6.1. Discontinuous Grammars

Discontinuous Grammars allow the rewriting rules to express a non-explicit sequence of symbols, terminal and/or nonterminal, with a special symbol **gap(X)**. A rule may refer to and re-order this non-explicit material. For example.

(6.1)      **verb, gap(X), nounphrase --> gap(X), verb, nounphrase.**

This rule expresses the repositioning of the **verb** to the right of the non-explicit sequence

**gap(X)** Of course, the non-explicit sequence may be empty (in which case no real repositioning would occur).

The power of DG rules stems both from the ability to reposition the non-explicit sequence(s) anywhere in the RHS of the rule, and from the fact that the non-explicit sequences may be any sequence of valid terminal or non-terminal symbols.

The opportunity appears to exist, then, to use DG rules to express linguistic phenomena that involve the repositioning of elements in a/the sequence of the strings of a language.

Let us briefly look at how this has been applied in the past. In [Dahl, Brown et al 86] we used the example (6.2) to show that that a complex noun phrase of the relative clause type could be analyzed with a DG

(6.2) **La casa que Julia construyó deslumbra.**  
*The house that Julia built glows.*

This was derived from the temporary intermediate structure of (6.3).

(6.3) **La casa rel\_mkr(X) [Julia construyó [la casa(X)]] deslumbra.**  
*The house rel\_mkr(X) [Julia built [the house(X)]] glows.*

Note that the linguistic analysis<sup>91</sup> here is that the NP [la casa(X)] undergoes movement to the relative-marker **rel\_mkr(X)** position and becomes a relative pronoun [que]. The DG rule used to express this movement was:

<sup>91</sup>A preferred linguistic analysis is that the complex NP has the phrase structure:

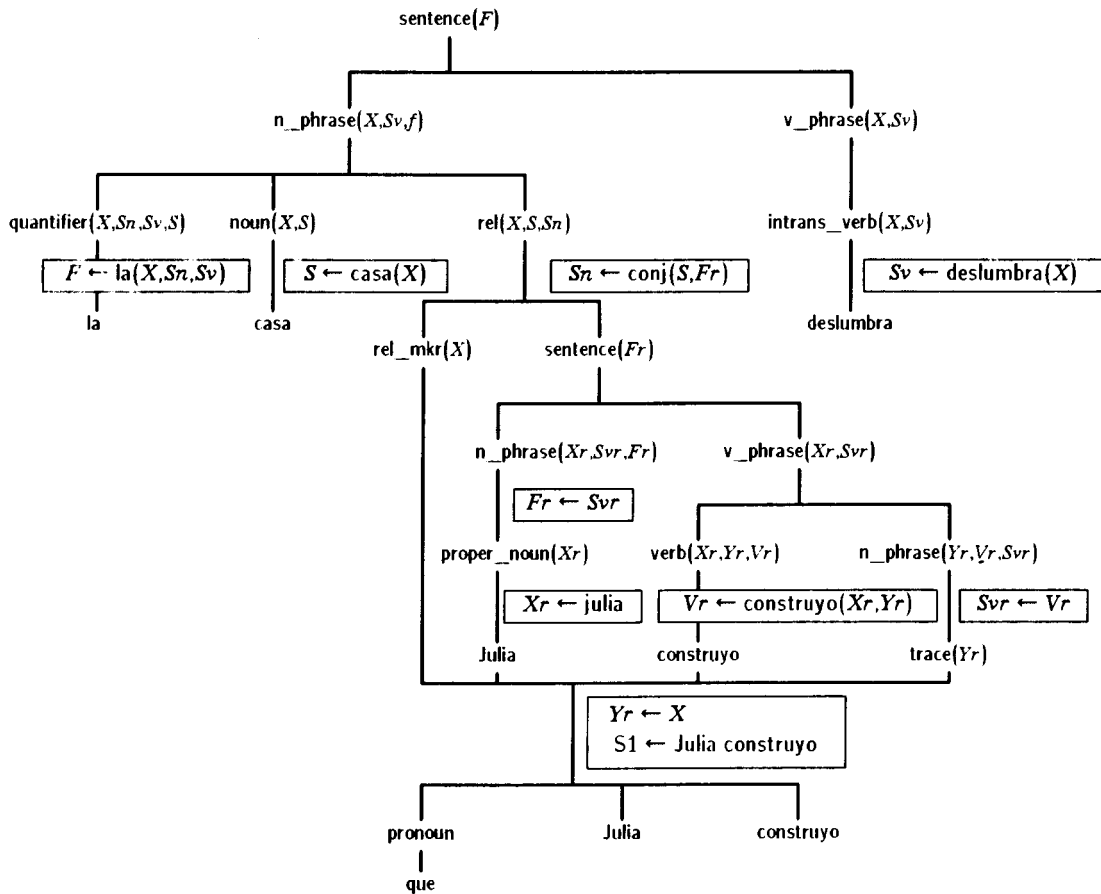
$$\left[ {}_{CP} \left[ {}_{CP} \left[ {}_{NP} \left[ {}_{NP} \text{ la casa} \right]_i \left[ {}_{CP} \left[ {}_{COMP} \text{ que} \right] \left[ \text{Julia construyó OP}_i \right] \right] \right] \right] \right] \text{ deslumbra } ]$$

The empty operator **OP** is co-indexed with its antecedent [la casa]. Of course, this analysis does not lend itself to the movement analysis discussed in [Dahl, Brown et al 86]. Current linguistic research in the area of complex NPs of the relative clause type centers around the question of *when* the co-indexation of **OP** and antecedent takes place

(6.4)  $rel\_mkr(X), gap(S), trace(X,Vp,F) \rightarrow rel\_pronoun, gap(S).$

The derivation graph for this analysis had the form given below (from [Dahl, Brown et al 86]).

(6.5)



Here we observe that the graph structure encountered presents a problem for the analysis of movement phenomena in GB Theory. It is currently believed [Chomsky 86] [Chomsky 82a] that the minimal phrasal tree structure involving the two constituents of a move- $\alpha$  phenomenon determine whether a constituent,  $\alpha$ , may in fact be allowed to move, or whether the resulting sentence would be ungrammatical. The principle of GB Theory involved here is the ECP (or Empty Category Principle) which states:

**Empty Category Principle: An empty category must be properly governed.**

Proper government involves either *antecedent government* or *lexical government*. Antecedent government, which concerns us here, means that the minimal tree structure involving the two constituents must have no *barriers* to government in the path from empty category (trace) to antecedent.

The fact that the derivational context of movement phenomena with DGs, in the old formulation, is a graph, and not a tree, means that it is impossible to apply the linguistic constraints on movement to the derivational context. In particular, it is not possible to check the antecedent government of a trace (an empty category) by its antecedent (the moved category) and thus the ECP cannot be applied. As many movement phenomena appear to involve antecedent government it is therefore impossible to verify whether a move  $\alpha$  transformation with DGs violates the ECP (and is consequently ungrammatical) for the antecedent government cases.

The graph structure of the derivation is inherent in DGs as originally formulated. The nature of the goal-reordering and graph structure is best illustrated by the following example from [Dahl 84]. For the language  $\{ a^n b^m c^n d^m \}$  the DG (6.6) was presented.

- (6.6.a)  $s \rightarrow as, bs, cs, ds.$
- (6.6.b)  $as, \text{gap}(G), cs \rightarrow [a], as, \text{gap}(G), [c], cs.$
- (6.6.c)  $as, \text{gap}(G), cs \rightarrow \text{gap}(G).$
- (6.6.d)  $bs, \text{gap}(G), ds \rightarrow [b], bs, \text{gap}(G), [d], ds.$
- (6.6.e)  $bs, \text{gap}(G), ds \rightarrow \text{gap}(G).$

These rules were compiled into the Prolog clauses (6.7) by the *synal* compiler.

(6.7.a)  $s(X0,X4) :- as(X0,X1),bs(X1,X2),cs(X2,X3),ds(X3,X4).$

(6.7.b)  $as(X0,X1) :-$   
 $connect(X0,a,T1),as(T1,T2),gap(G,T2,T5),$   
 $connect(T5,c,T6),cs(T6,T7),gap(G,X1,T9),$   
 $cs(T9,T7).$

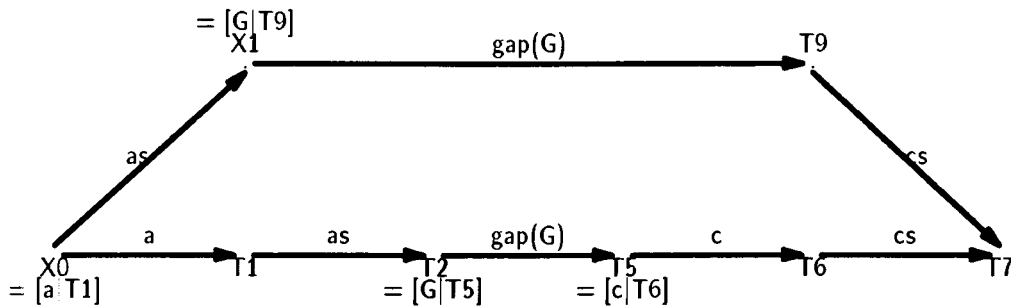
(6.7.c)  $as(X0,X1) :-$   
 $gap(G,X0,T4),gap(G,X1,T5),cs(T5,T4).$

(6.7.d)  $bs(X0,X1) :-$   
 $connect(X0,b,T1),bs(T1,T2),gap(G,T2,T5),$   
 $connect(T5,d,T6),ds(T6,T7),gap(G,X1,T9),$   
 $ds(T9,T7).$

(6.7.e)  $bs(X0,X1) :-$   
 $gap(G,X0,T4),gap(G,X1,T5),ds(T5,T4).$

An examination of the first rule for  $as$  gives the following graphical display<sup>92</sup> of the rule, where nodes represent phrase boundaries and labelled arcs represent non-terminals.

(6.8)



The interpretation makes use of the two extra input/output list arguments of DCGs in the construction of the derivation graph through the re-ordering of the goals. Note how the arguments of the goals in the compiled rules correspond to the structure of the graph.

<sup>92</sup>In the notation of [Dahl and Abramson 84].

It is the DCG input/output list arguments that are used to construct the derivational context graph. While we use these list arguments in the new formulation as discussed later they are not crucial to the formalism, and in fact in [Dahl and Brown 86] we used the formalism without these arguments to program applications in non-linguistic areas of computing such as parallel-processing.

The non-explicit material noted with `gap(X)` in a DG rule was interpreted in its compiled form by the rules (6.9) given in the `synal` compiler.

(6.9.a) `gap([Word|List]) --> [Word], gap(List).`

(6.9.b) `gap([]) --> [].`

The symbol `gap(X)` can be thought of as a kind of append operation. In compiled form these gap rules become the rules of (6.10).

(6.10.a) `gap([Word|List],X0,X1) :- connect(X0,Word,Temp),  
gap(List,Temp,X1).`

(6.10.b) `gap([],X0,X0).`

This leads to the second of our problems with DGs in the original formulation in linguistic applications. While DGs can and have been used to express extraposition, free word order, and coordination in NL analysis [Dahl 84] in the original formulation, it turns out that they cannot be used for NL generation (synthesis of the output strings).

The reason is quite straightforward and I shall illustrate it here with a simple example. Consider the fragment of a toy DCG grammar (6.11) constructed along the lines of the the generation grammar for Spanish of this thesis.

- (6.11.a)  $xp(xp, xp(e), empty) \rightarrow [e]$ .
- (6.11.b)  $xp(ip, ip(PT1, PT2), [+pass]) \rightarrow spec(ip, PT1, [+pass]),$   
 $xbar(ip, PT2)$ .
- (6.11.c)  $spec(ip, PT, [+pass]) \rightarrow xp(xp, PT, empty)$ .
- (6.11.d)  $xbar(ip, ibar(PT1, PT2)) \rightarrow x(ip, PT1), compl(ip, PT2)$ .
- (6.11.e)  $x(ip, infl(tns, agr)) \rightarrow [infl(tns, agr)]$ .
- (6.11.f)  $compl(ip, PT) \rightarrow xp(vp, PT, F)$ .
- (6.11.g)  $xp(vp, vp(PT1, PT2), [+tns]) \rightarrow spec(vp, PT1, [+tns]),$   
 $xbar(vp, PT2)$ .
- (6.11.h)  $spec(vp, [], [+tns]) \rightarrow []$ .
- (6.11.i)  $xbar(vp, vbar(PT1, PT2)) \rightarrow x(vp, PT1), compl(vp, PT2)$ .
- (6.11.j)  $x(vp, verb(see)) \rightarrow [see]$ .
- (6.11.k)  $compl(vp, PT) \rightarrow xp(np, PT, F)$ .
- (6.11.l)  $xp(np, np(PT1, PT2), def) \rightarrow spec(np, PT1, def),$   
 $xbar(np, PT2)$ .
- (6.11.m)  $spec(np, det(the), def) \rightarrow [the]$ .
- (6.11.n)  $xbar(np, nbar(PT1, PT2)) \rightarrow x(np, PT1), compl(np, PT2)$ .
- (6.11.o)  $x(np, noun(house)) \rightarrow [house]$ .
- (6.11.p)  $compl(np, []) \rightarrow []$ .
- (6.11.q) done.

When this grammar is asked to parse (6.12) an input string, or to generate (6.13) an input string it does so quite happily.

(6.12)

C-Prolog version 1.5

| ?- [trial23].

trial23 consulted 2360 bytes 0.45 sec.

yes

| ?-  $xp(ip, PT, F, [e, infl(tns, agr), see, the, house], [])$ .

$PT = ip(xp(e), ibar(infl(tns, agr), vp([], vbar(verb(see),$   
 $np(det(the), nbar(noun(house), []))))))$

$F = [+pass]$  ;

no

(6.13)

| ?- xp(ip,PT,F,Phrase,[]).

PT = ip(xp(e),ibar(infl(tns,agr),vp([],vbar(verb(see),  
np(det(the),nbar(noun(house),[])))))

F = [+pass]

Phrase = [e,infl(tns,agr),see,the,house] ;

no

Now we convert the grammar into a DG by Replacing rule (6.11.a) with (6.14.a). and adding the rule (6.14.b)

(6.14.a) xp(xp,PT1,empty), gap(X), xp(np,PT1,Feat) -->  
xp(np,PT1,Feat), gap(X), xp(np,np(t),trace).

(6.14.b) xp(np,np(t),trace) --> [t].

When this DG is compiled by the `synal` preprocessor for DGs and run to analyse (6.15) a string it works fine. The internal representation of a parse tree (PT) is created. the input string ([the,house,X,see,t]) is analysed. and the feature variable (Feat) is bound to [+pass].



**(6.15)****C-Prolog version 1.5****| ?- [synal].****synal consulted 1800 bytes 0.333333 sec.****yes****| ?- sa(trial21).****yes****| ?- xp(ip,PT,Feat,[the,house,X,see,t],[]).****PT = ip(np(det(the),nbar(noun(house),[])),ibar(infl(tns,agr),  
vp([],vbar(verb(see),np(det(the),nbar(noun(house),[])))))****Feat = [+pass]****X = infl(tns,agr) ;****no**

However, when the DG is asked to generate (6.16) a string (the unbound variable **Phrase**), an infinite recursive loop occurs.

(6.16)

C-Prolog version 1.5

| ?- [syal].

syal consulted 1800 bytes 0.333333 sec.

yes

| ?- [pt].

pt consulted 604 bytes 0.0500003 sec.

yes

| ?- listing(parsetree).

```

parsetree(ip(np(det(the),nbar(noun(house),'[]')),
             ibar(infl(tns,agr),
                  vp('[]',
                     vbar(verb(see),
                           np(det(the),nbar(noun(house),'[]'))
                           )))))

```

yes

| ? sa(trial21).

yes

| ? parsetree(PT),xp(ip.PT,[+pass],Phrase,[]).

! Out of local stack during execution

[ execution aborted ]

A trace of execution (6.17) reveals the reason why. The call to **gap** in the generation mode will inevitably result in a call with three unbound variables.

It may be seen in call (12) of the trace that the call to **gap(\_38,\_37,\_39)** involves three unbound variables. Such a call will always go into a recursive loop (involving the built-in predicate **c (connect)**) as it attempts to append one unbound variable representing a list to another. Intuitively, the attempt to generate a *skipped* sequence of terminals of unknown length is our downfall. The grammar continues to append skipped sequences to the string being synthesised because there is no mechanism to stop it. There is no input string of terminals (as in analysis). Thus there is no way to check whether the next terminal should be skipped (ie. continue the 'gap') or not (ie. terminate the gap).

## (6.17)

| ?- trace.

yes

```

| ?- parsetree(PT).xp(ip,PT,[+pass],Phrase,[]).
(1) 1 Call: parsetree(_0) ?
(1) 1 Exit: parsetree(ip(np(det(the),nbar(noun(house),[])),
      ibar(infl(tns,agr),vp([],vbar(verb(see),
      np(det(the),nbar(noun(house),[])))))
(2) 1 Call: xp(ip,ip(np(det(the),nbar(noun(house),[])),
      ibar(infl(tns,agr),vp([],vbar(verb(see),
      np(det(the),nbar(noun(house),[])))))],[+pass],_11,[]) ?
(3) 2 Call: spec(ip,np(det(the),nbar(noun(house),[])),[+pass],_11,_65645) ?
(4) 3 Call: xp(xp,np(det(the),nbar(noun(house),[])),empty,_11,_65645) ?
(5) 4 Call: xp(np,np(det(the),nbar(noun(house),[])),_36,_11,_37) ?
(6) 5 Call: spec(np,det(the),def,_11,_65677) ?
(7) 6 Call: c(_11,the,_65677) ?
(7) 6 Exit: c([the|_58],the,_58)
(6) 5 Exit: spec(np,det(the),def,[the|_58],_58)
(8) 5 Call: xbar(np,nbar(noun(house),[]),_58,_37) ?
(9) 6 Call: x(np,noun(house),_58,_65706) ?
(10) 7 Call: c(_58,house,_65706) ?
(10) 7 Exit: c([house|_66],house,_66)
(9) 6 Exit: x(np,noun(house),[house|_66],_66)
(11) 6 Call: compl(np,[],_66,_37) ?
(11) 6 Exit: compl(np,[],_37,_37)
(8) 5 Exit: xbar(np,nbar(noun(house),[]),[house|_37],_37)
(5) 4 Exit: xp(np,np(det(the),nbar(noun(house),[])),def,[the,house|_37],_37)
(12) 4 Call: gap(_38,_37,_39) ?
(13) 5 Call: c(_37,_79,_65743) ?
(13) 5 Exit: c([_79|_84],_79,_84)
(14) 5 Call: gap(_80,_84,_39) ?
(15) 6 Call: c(_84,_87,_65762) ?
(15) 6 Exit: c([_87|_92],_87,_92)
(16) 6 Call: gap(_88,_92,_39) ? ^CAction (h for help): a

```

[ execution aborted ]

The conclusion we reached was that a new interpretation of DG rules would be necessary. In the next section I discuss the new formulation we presented in [Dahl, Brown et al 86] to solve these two problems: the graph structure of the derivation, and the recursive loops for generation. Patrick Saint-Dizier undertook the implementation of these ideas [Saint-Dizier 87] and noted that, in his implementation, generation would work.

## 6.2. A New Formulation for Discontinuous Grammars

We have seen in the last section the problems involving the linguistic applications of DGs. Namely: a flattening of the derivational context implied by the graphical nature of the derivational formalism, a concomitant absence of a true derivational tree structure necessary to GB Theory, and recursive loops involving non-explicit material associated with the  $\text{gap}(X)$  rules.

A new set of conventions was proposed in [Dahl, Brown et al 86]. These were specifically aimed at providing an interpretation of rules in NL applications based on GB Theory:

- **Convention 1:** Each rule with more than one left hand side symbol must have an equal number of symbols on each side of the rewriting arrow.
- **Convention 2:** A **skip** symbol or a terminal symbol can only be rewritten into its corresponding position.
- **Convention 3:** The application of a rule

$$a_1, a_2, \dots, a_n \rightarrow a_1', a_2', \dots, a_n'.$$

is interpreted as the simultaneous application of the rules

$$\begin{array}{l} a_1 \rightarrow a_1' \\ a_2 \rightarrow a_2' \\ \vdots \\ a_n \rightarrow a_n' \end{array}$$

Saint-Dizier has pointed out<sup>93</sup> that **Convention 1** is overly restrictive in NL applications.

The conventions are formally abandoned here and replaced with the following:

- **Convention 1a:** Each rule with more than one element in the left hand side must have an equal number of skips on the left hand side and the right hand side.

---

<sup>93</sup>Personal communication. See also. [Saint-Dizier 87].

- **Convention 2a:** Each rule with more than one element on the left hand side must have each non-terminal element on that side separated by **skip** symbols. Each of these non-terminal elements must have a corresponding string of terminals, non-terminals, and Prolog calls on the right hand side.
- **Convention 3a:** The application of the rule

$$\begin{array}{l}
 \alpha_1, \text{ skip}, \alpha_2, \text{ skip}, \alpha_3, \dots, \text{ skip}, \alpha_n \text{ -->} \\
 \quad \alpha_{11}, \alpha_{12}, \dots, \alpha_{1w}, \\
 \quad \text{skip}, \\
 \quad \alpha_{21}, \alpha_{22}, \dots, \alpha_{2x}, \\
 \quad \text{skip}, \\
 \quad \alpha_{31}, \alpha_{32}, \dots, \alpha_{3y}, \\
 \quad \text{skip}, \\
 \quad \dots \\
 \quad \alpha_{n1}, \alpha_{n2}, \dots, \alpha_{nz}.
 \end{array}$$

is interpreted as the simultaneous application of the rules

$$\begin{array}{l}
 \alpha_1 \text{ -->} \alpha_{11}, \alpha_{12}, \dots, \alpha_{1w}, \\
 \alpha_2 \text{ -->} \alpha_{21}, \alpha_{22}, \dots, \alpha_{2x}, \\
 \alpha_3 \text{ -->} \alpha_{31}, \alpha_{32}, \dots, \alpha_{3y}, \\
 \\
 \alpha_n \text{ -->} \alpha_{n1}, \alpha_{n2}, \dots, \alpha_{nz}.
 \end{array}$$

where the simultaneously applied rules share variables through unification.

The effect is to provide for the simultaneous application of DCG-type rules. Thus, rules such as (6.18) are the basic paradigm.

$$(6.18) \quad a, \text{ skip}, b \text{ -->} a', \text{ skip}, b'.$$

Rules such as (6.19) are possible and are interpreted as the simultaneous application of two rules (6.20).

$$(6.19) \quad a, \text{ skip}, b \text{ -->} a_1', a_2', \{a_3'\}, \text{ skip}, [b_1'], b_2'.$$

$$(6.20.a) \quad a \text{ -->} a_1', a_2', \{a_3'\}.$$

$$(6.20.b) \quad b \text{ -->} [b_1'], b_2'.$$

The conventions 1a, 2a, and 3a have the potential for solving the problems of a flattened derivational context and lack of a derivational tree upon which constraints such as *barriers* to government and subjacency may be applied. An implementation of these conventions needs also to avoid the infinite recursive loop problem for NL generation pointed out in connection with the DG formalism. We proceed to such an implementation now.

An extension to this new formulation by Saint-Dizier<sup>94</sup> has introduced adjunction operations allowing the grammar writer to express adjunction as well as substitution transformations.

### 6.3. An Implementation of a Discontinuous Grammar Compiler under the New Formulation

The new DG formulation, as we have seen, was developed from the need to deal with GB theory in the context of NL generation applications [Dahl, Brown et al 86]. Given that this formulation is simpler (more regular) than the previous one, it allows, as one would expect, for simpler implementations. Patrick Saint-Dizier undertook a first implementation of the new formulation. His implementation of the simultaneous application of rules is discussed in section 6.3.2.

The *synapse3* compiler found in appendix Z was developed for this dissertation in order to test the computational ideas of the preceding section and the transformational ideas of Chapter 7. DG rules are compiled into Horn clauses suitable for interpretation by the

---

<sup>94</sup>In the forthcoming paper *Contextual Discontinuous Grammars* (to be read at the Second International Workshop on Natural Language Understanding and Logic Programming, Simon Fraser University, August 1987) Saint Dizier uses a different notation for expressing the transformations.

**c-prolog** interpreter Included is the ability to apply the conventions {1a.2a.3a} and give the effect of simultaneous applications of rules. Also included is the automatic buildup of derivational context reported by [Dahl and St-Dizier 86] so that constraints based on ancestor relations in the derivational context may be applied.

### 6.3.1. Adding the derivational context

In [Dahl and St-Dizier 86] a derivational context was used to provide a basis for the application of constraints on the old formulation of DGs. The following discussion shows how such a derivational context is added to Prolog rules. Dahl and Saint-Dizier called the resulting grammar a Constrained Discontinuous Grammar(CDG). They report that constraints (such as **Subjacency**) on transformations are easily added to a CDG. Because the **synapse3** compiler adds a derivational context in the same manner, I call the resulting grammar a CDG

A rule (6.21) for the predicate **a**, if it succeeds, will have called the procedures **b**, **c**, and **d**

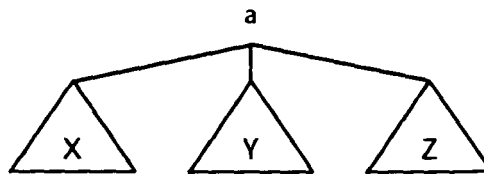
$$(6.21) \quad a(A1,A2) \text{ --> } b(B1), \\ \qquad \qquad \qquad c(C1,C2,C3), \\ \qquad \qquad \qquad d(D1,D2).$$

The derivational context of a Prolog rule may be thought of as a tree whose root is the head of the rule, whose nodes are the goals and subgoals of the resolution proof procedure and whose leaves are the assertions verified in the database.

The derivational context of (6.21) is thus a tree of the form (6.22).

This tree, in the form of a Prolog functor, is **a(X,Y,Z)**. Here, {**X,Y,Z**} are unbound variables representing the (unknown) derivational contexts of the subgoals. The number of unbound variables corresponds to the number of subgoals in the body of the rule.

(6.22)



We can add this functor to the rule as an extra argument of the head of the rule, generating the new head (6.23).

(6.23)  $a(a(X,Y,Z),A1,A2)$ 

and providing an extra argument in each of the subgoal predicates as in (6.24).

(6.24.a)  $b(b(X),B1)$ (6.24.b)  $c(c(Y),C1,C2,C3)$ (6.24.c)  $d(d(Z),D1,D2)$ 

The resulting rule with derivational context is (6.25).

(6.25)  $a((a(X,Y,Z),A1,A2) \rightarrow b(X,B1),$   
 $c(Y,C1,C2,C3),$   
 $d(Z,D1,D2).$

It turns out to be convenient to have another extra argument representing the whole derivational context for those cases when  $a$  is not the root goal of a proof, but only a subgoal. This extra argument ( $DC$ , for Derivational Context) is added to all predicates of a rule in the head and the body. This, of course, must be done to each and every rule for  $a$ . Similarly, every call to  $a$  in the database must now include an extra argument. The resulting rule is, then, (6.26).

(6.26)  $a(DC,(a(X,Y,Z),A1,A2) \rightarrow b(DC,X,B1),$   
 $c(DC,Y,C1,C2,C3),$   
 $d(DC,Z,D1,D2).$



### 6.3.2. Simultaneous application of rules

In the **synal** compiler for the original formulation of DG rule interpretation, a list of unresolved goals was used to store those goals skipped over in the derivation.

For the new formulation for DGs, we want to simulate the simultaneous application of rules to a derivational context.

The interpretation becomes: Where two rules were the interpretation of a single DG rule, then the first rule is considered to succeed if and only if the other rule also succeeds. To maintain this interpretation, the information that some goals remain to be resolved in the future is stored. However, unlike Dahl's method for DGs in the original formulation the unresolved **goals** are not stored. Instead, using the implementation of Saint-Dizier, the head and the body of the **rule(s)** still to be applied are stored.

This does not accomplish the same thing as a **prolog cut**. With a cut, no backtracking to the parent goal is permitted. With our new formulation of DG rules, we want to permit backtracking where goals do not succeed.

The **freeze** predicate of **Prolog II** permits the suspension of the execution of goals when a variable is unbound. The execution resumes when the variable is bound by the successful unification of the variable to a non-variable during the execution of another goal. Our formulation bears more resemblance to this methodology. Both allow backtracking when goals fail. The difference, again, is that we suspend execution of a specific rule (among perhaps several that might apply to the head of the suspended rule.)

Recall that the rule (6.27) was to be interpreted as the simultaneous application of two rules (6.28) with shared variables.

**(6.27) a, skip, b --> ap, skip, bp.**

**(6.28.a) a --> ap.**

**(6.28.a) b --> bp.**

This is accomplished by storing the second rule as a rule to be applied in the future if the first rule is to succeed as in (6.29). To store the heads and bodies of rules for future execution the pair (head,body) is concatenated to an input list (**InRules**) of stored rules, to give an output list (**OutRules**) of stored rules.

**(6.29) a --> ap, conc(InRules, [store(b,bp)], OutRules).**

The second rule must remove the stored (head,body) pair from the input list of stored rules as in (6.30). If the second rule itself succeeds, the requirements for application of the first rule are completed. As shown, this is done before the predicates of the second rule are applied to avoid loops.

**(6.30) b --> remove(store(b,bp),InRules,OutRules), bp.**

This seems straightforward enough, however, in order to share variables, the rules must have the input list (**InRules**) and output list (**OutRules**) added explicitly to the argument lists for each predicate.

This is done by adding these arguments automatically during compilation. The writer of a CDG rule of this type is relieved of the need to worry about maintaining the list of stored rules. Because two prolog rules cannot actually share variables<sup>95</sup> except through the process of passing variables as parameters to subgoals, the methodology of adding these lists explicitly accomplishes the sharing.

---

<sup>95</sup>All variables in prolog are universally quantified and local to the individual prolog rule.

(6.31.a) **a(InRules1,OutRules1):-**  
     **ap(InRules1,TempRules1),**  
     **conc(TempRules1,[store(b,bp)],OutRules1).**

(6.31.b) **b(InRules2,OutRules2):-**  
     **remove(store(b,bp),InRules2,TempRules2),**  
     **bp(TempRules2,OutRules2).**

In summary, the four tasks of the compiler are to:

- Break the CDG rule into two rules (or more) rules.
- Add the derivational context arguments.
- Add the stored-rule-lists arguments, and (concatenation and removal) predicates.
- Add the DCG list arguments.

When all this is accomplished, the simple rule (6.27) becomes a pair of rules with a complex argument structure (6.32). The derivational context arguments are **a(X)**, **b(Y)**, **X**, **Y**, and **DC**, added at the beginning of the argument list for each predicate. The stored rules (containing shared variables) are in the **InRules<sub>i</sub>**, **TempRules<sub>j</sub>**, and **OutRules<sub>k</sub>** arguments added to the end of the argument list for each predicate. Finally, the internal DCG preprocessor of the c-prolog interpreter is allowed to add the list variables **InList<sub>i</sub>** and **OutList<sub>j</sub>** to the ends of the argument lists.

(6.32.a) **a(a(X),DC1,InRules1,OutRules1,InList1,OutList1):-**  
     **ap(X,DC1,InRules1,TempRules1,InList1,OutList1),**  
     **conc(TempRules1,[store(b,bp)],OutRules1).**

(6.32.b) **b(b(Y),DC2,InRules2,OutRules2,InList2,OutList2):-**  
     **remove(store(b,bp),InRules2,TempRules2),**  
     **bp(Y,DC2,TempRules2,OutRules2,InList2,OutList2).**

## 6.4. Summary

This Chapter has discussed some problems with linguistic applications of DGs in the original formulation; the flattening of the derivational tree into a graph-like derivation, and recursive looping when used for generation of strings. The linguistic motivation for a true tree as a derivational context was introduced in a discussion of antecedent government.

A new formulation of DGs was introduced. The motivation was to avoid recursive loops and to have derivational trees available for constraint application. These requirements suggested an interpretation of DG rules as the simultaneous application of two or more prolog rules derived from one DG rule, with the additional feature of variables shared through unification.

The action of a compiler of DG rules into c-prolog rules was discussed. (The compiler itself is given in Appendix B.) The role of the compiler is to relieve the grammar writer of the task of adding the extra arguments and predicates to the grammar rules.

In the next Chapter we shall see some examples of specific DG rules for transformations that have been tested with the compiler.

## Chapter 7

### Transformations and Spanish S-structure

This chapter gives details of the transformations used in the grammar in order to generate the S-structure of Spanish sentences with clitics.

The Projection Principle states that: *Representations at every syntactic level are projected from the Lexicon.* That is, the subcategorisation properties and thematic relations of lexical items must be observed at every level. Thus in both D-structure (before transformations) and at S-structure (after transformations) the subcategorisation properties and  $\Theta$ -roles must be observed.

One of the tasks of the grammar writer is to ensure that subcategorisation properties, Case Theory, and Theta Theory are observed in the rules of the grammar. We shall see in the Move-XP transformation given in the next section that careful attention must be paid to both Case and  $\Theta$ -roles.

#### 7.1. Move-XP Transformations

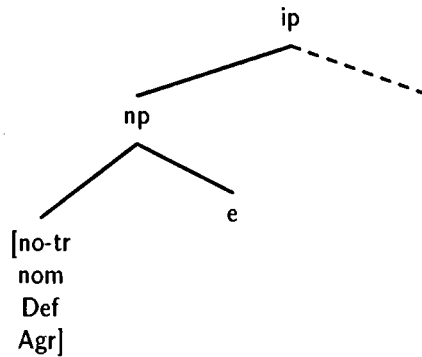
XPs may move to an empty XP position subject to the constraints of the grammar.

As an example of this kind of transformation, consider the true passive construction. In the generation of true passives, the SPEC of IP is generated in D-structure as an empty NP position by rule (7.1). This empty XP position is assigned no  $\Theta$ -role (**no-tr**) and nominative case (**nom**) by the rule.

(7.1)

$xp(np, PT, G, [no-tr, nom, Def, Agr, Overt, Animate]) \rightarrow$   
 $[e],$   
 $\{PT = ..[xp, [no-tr, nom, Def, Agr], e]\}.$

(7.2)



A CDG rule that moves an XP to this position will begin with the predicate  $xp(np, PT, G, [no-tr, nom, Def, Agr, Overt, Animate])$  found at the head of (7.1). Such a rule is the Move-NP transformation introduced next.

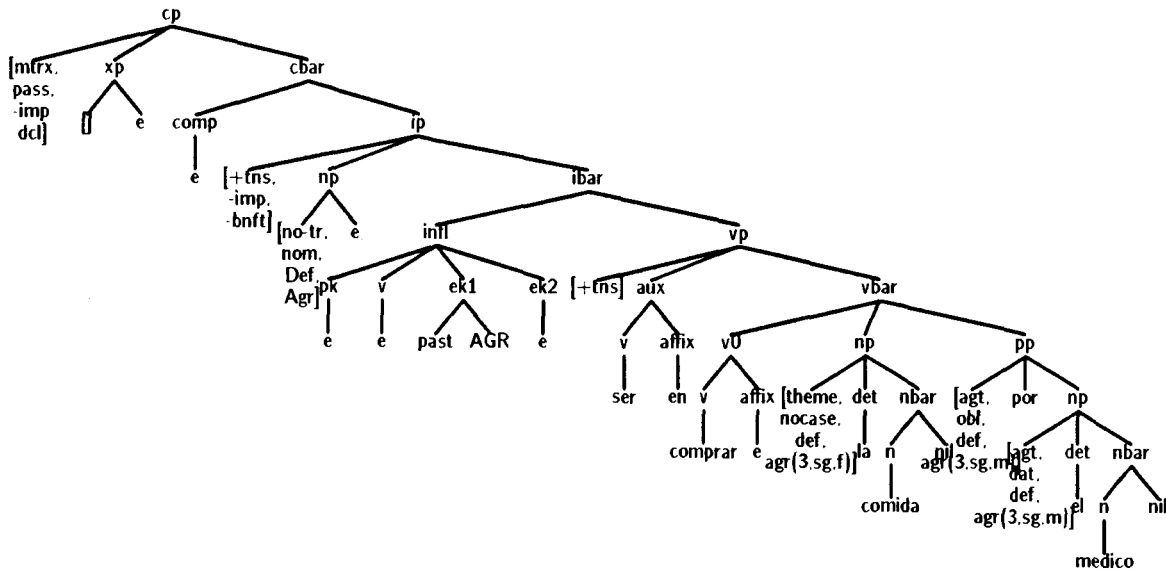
### 7.1.1. A Move-NP transformation

Recall that the generation of true passive D-structure produced an NP without Case in the verb complement position. Such an NP violates the Case Filter<sup>96</sup>.

The complete output example (7.3) gives a complete D-structure (with both an empty NP and an NP without Case) for a true passive sentence. As it is, this sentence violates the Case Filter. If the NP were to move to a position where it is assigned Case, however Case theory would not rule the resulting sentence ungrammatical.

<sup>96</sup>The Case Filter states that every overt NP must be assigned Case.

(7.3) **La comida fue comprada por el médico.** (D-structure)  
*The food was bought by the doctor.*



What is needed, then is a transformation that moves the NP without Case to an empty XP position where Case may be assigned. In the process, a trace [t] of the moved NP must be left in the position vacated by the NP. This is an example of an active constraint on generation. The CDG rule that does this is (7.4). This rule ensures that Case is assigned to an NP without Case, and that a the syntactic chain from moved NP to trace has only one  $\Theta$ -role.

(7.4)

$$\begin{aligned} &xp(np, PT1, G, [no-tr, nom, Def1, Agr1, Ovr1, Animate1]), \\ & \quad skip, \\ & \quad xp(np, PT2, G, [TR, nocase, Def2, Agr2, Ovr2, Animate2]) \\ \rightarrow & \\ & xp(np, PT1, G, [TR, nom, Def2, Agr2, Ovr2, Animate2]), \\ & \quad skip, \\ & \quad xp(nptrace, PT2, G, [TR, trace, Def2, Agr2, Ovr2, Animate2]). \end{aligned}$$

Of course, as we saw in Chapter 6, this is equivalent to the simultaneous application of the two rules in (7.5) where the two rules share variables through unification.

(7.5)

xp(np,PT1,G,[no-tr,nom,Def1,Agr1,Ovrt1,Animate1]),

-->

xp(np,PT1,G,[TR,nom,Def2,Agr2,Ovrt2,Animate2]),

xp(np,PT2,G,[TR,nocase,Def2,Agr2,Ovrt2,Animate2])

-->

xp(nptrace,PT2,G,[TR,trace,Def2,Agr2,Ovrt2,Animate2]).

An additional rule (7.6) to generate the trace itself is required.

(7.6)

xp(nptrace,PT,G,[theme,trace,Def,Agr,Ovrt,Animate]) -->

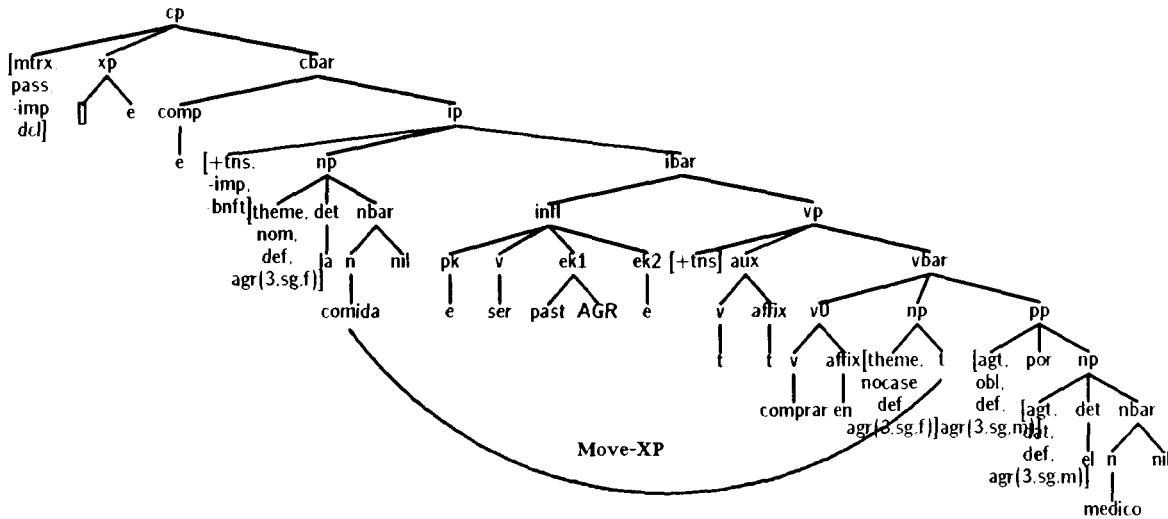
!, [t],

{PT=..[np,[theme,nocase,Def,Agr],t]}.

The S-structure resulting from this (and other) transformations is (7.7). (The other transformations that apply are treated in the following sections of this Chapter.)

(7.7) La comida fue comprada por el médico. (S-structure)

*The food was bought by the doctor.*





## 7.2. Move-X Transformations

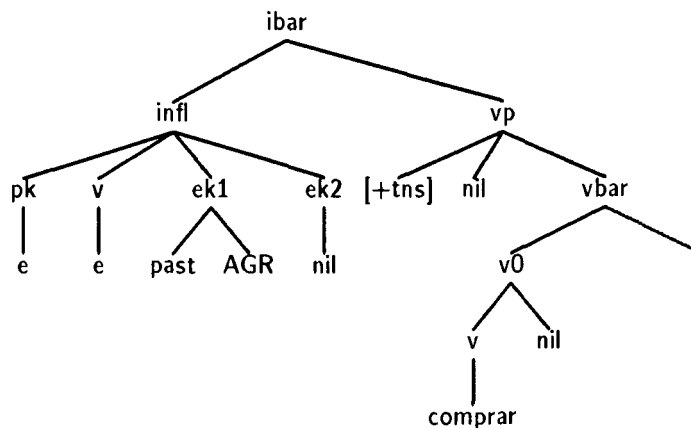
Heads-of-XPs may move to an empty Head-of-XP position. Recall that a Head-of-XP is the X of the XP structure. The Xs are therefore lexical items such as verbs, inflections, et cetera.

The transformations that move Heads to Head-of-XP positions are relatively unexplored in the literature. One such movement has been suggested in Barriers [Chomsky 86]: the movement of a verb from Head-of-VP position to INFL (Head-of-IP) to become an inflected verb. This transformation is discussed next.

### 7.2.1. The Verb to INFL transformation

The generation of D-structure gave an empty X-position in INFL and left the verb in VP as illustrated in the fragment (7.8). The empty X-position in INFL (e under v in infl) was generated with rule (7.9). (An auxiliary verb might also be present in the D-structure VP. This situation will be treated a little later when Affix-Hopping is treated.)

(7.8) D-structure fragment



(7.9)

$x(\text{infl}, \text{PT}, \text{G}, \text{F}) \rightarrow [e], \{ \text{PT} = ..[\text{infl}, e] \}$ .

I have expressed the Move-X transformation of verb to INFL as a single CDG rule (7.10). The rule moves the verb to the inflectional structure of INFL, leaving a trace [t] in the Head-of-VP position. It is quite general in nature, with only the categories involved named specifically in the rule.

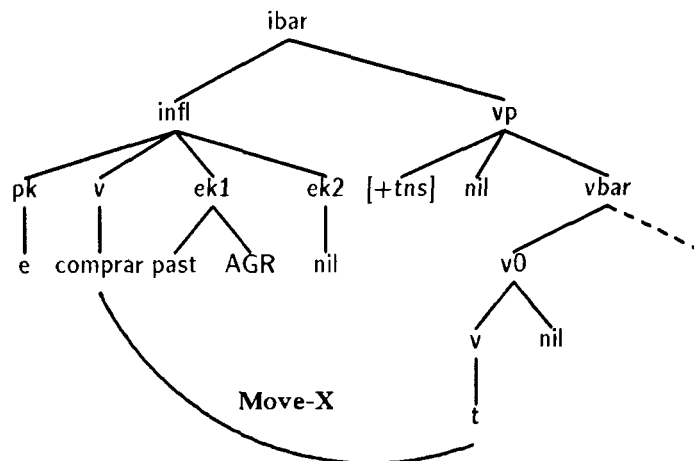
(7.10)  
 $x(\text{infl}, \text{PT1}, \text{G}, \_)$ , skip,  $x(\text{vp}, \text{PT2}, \text{G}, \text{W}) \rightarrow$   
 $x(\text{vp}, \text{PT1}, \text{G}, \text{W})$ , skip,  $x(\text{vtrace}, \text{PT2}, \text{G}, \text{W})$ .

As before, a rule is used to insert the trace in the internal structure.

(7.11)  
 $x(\text{vtrace}, \text{PT}, \text{G}, \_)$   $\rightarrow$  [t], {PT=..[v,t]}.

For simple tenses the matrix verb is moved to INFL as in (7.12) by the rule (7.11). Because of its generality, this rule will move an auxiliary (instead of the main verb) to INFL if that auxiliary is encountered in the structure before the main verb. The movement of the auxiliary verb would leave a stranded affix unless a rule for affix-hopping is provided (Just such a rule is discussed in the next section.)

(7.12) (fragment of S-structure)



### 7.3. Minor Movement Rules

Minor Movement Rules are those transformations that do not fall under the categories of either Move-X or Move-XP.

#### 7.3.1. The Affix-Hopping transformation

Affix-Hopping is the movement of an affix associated with the voice, tense and aspect systems of the VP to a position adjoined to a verb. If an auxiliary verb is Base generated, an accompanying affix is also generated. Consider the example of the true passive construction. In Spanish this construction requires the insertion of the auxiliary verb *ser* and the past participial affix *en*.

It is assumed that the auxiliary is inserted in the SPEC-of-VP position in the generation of D structure as seen in (7 13). Affix-Hopping will then move the affix to adjoin to the verb<sup>97</sup> in Head-of-VP position as indicated. (The Move-X transformation will move the auxiliary verb, not the main verb, to INFL.)

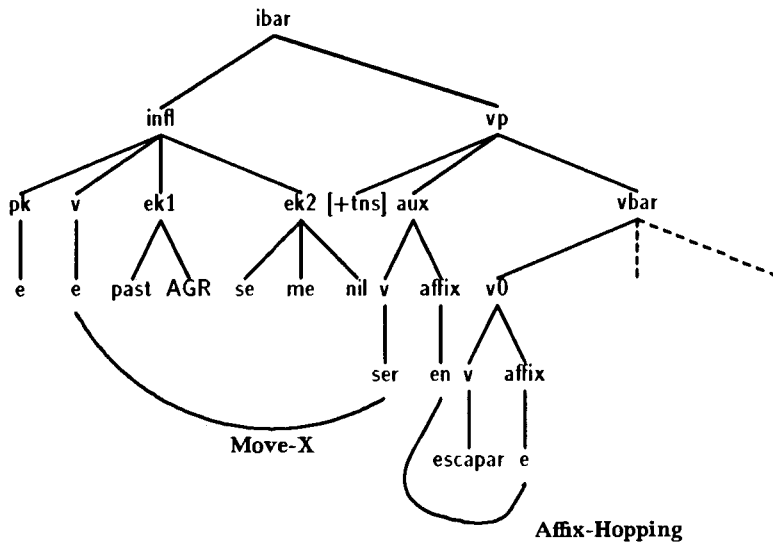
When the auxiliary is moved from this position to INFL, the stranded affix must also move to become the affix of a verb. This is the well known affix-hopping phenomenon of [Chomsky 77].

---

<sup>97</sup>Other possible analyses of the derivation of the participle exist. One is that no affix hopping occurs. Instead, a feature system replaces the affix. This feature system licenses the inflection of the verb at PF.

Another is that it is not the affix that moves, but the verb. The verb cannot occupy the position vacated by the auxiliary, because that position is occupied by a trace. Thus, the verb must move to adjoin to the trace. This second possibility has the advantage that proper government of the trace of the moved verb by its antecedent is achieved. The analysis I have used in the generation program does not achieve proper government of the trace of the affix. The affix-hopping analysis was chosen because adjunction movement is not expressible in the CDG rules I have presented. Currently, Saint-Dizier is developing adjunction mechanisms for CDGs.

## (7.13) (D-structure fragment)



The CDG rule for affix-hopping is given in (7.14). This represents a movement to the right, from the D-structure position to adjoin to the verb.

## (7.14)

**affix(pass,PT1,W), skip, affix(vp,PT2,nil) -->**  
**affix(trace,PT1,W), skip, affix(pass,PT2,W).**

The interpretation of this rule as the simultaneous application of the two rules (7.16) is as follows (7.16.a) says, declaratively, that a passive affix with internal structure **PT1** and form **W** is rewritten as a trace with the same internal structure and form. Simultaneously, application of (7.16.b) says that the empty affix position in VP with internal structure **PT2** is rewritten as a passive affix.

(7.16.a) **affix(pass,PT1,W) --> affix(trace,PT1,W).**

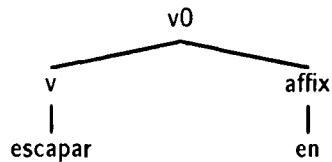
(7.16.b) **affix(vp,PT2,e) --> affix(pass,PT2,W).**

Application of rule (7.17) actually inserts the trace.

(7.17) **affix(trace,PT,W) --> [t], {PT=..[affix,t]}.**

The structure (7.18) resulting from affix-hopping is interpreted through the morphological and phonological components of grammar. For example, the verb-plus-affix construction becomes a participial form of the verb (7.19).

(7.18)



This dissertation does not treat the morphology or phonology of Spanish. (7.19) is given only to make the interpretation of S-structure easier.

(7.19)

*verb* + *affix* becomes *participle*

**escapar** + **en** becomes **escapado**

The S-structure for a complete sentence resulting from this (and other) transformations was given in (7.7).

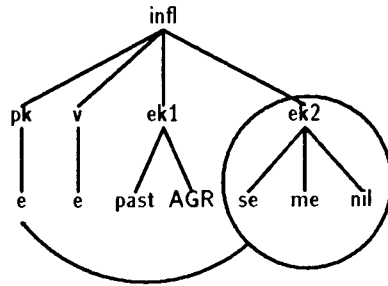
Affix-hopping is an example of a minor movement rule that moves an inflectional element (the affix) to the right. Now we examine a mechanism that moves an inflectional element to the left.

### 7.3.2. Clitic Hopping

Clitic Hopping as defined in Chapter 4 is the movement (see (7.20)) of clitic inflections from ENCL2 position (**ek2**) to the empty PROCL position (**pk**) under the control of tense. If the IP has the [TNS] feature unified with [+tns], the clitic(s) in ENCL2 may optionally move to PROCL. Normally, in a tensed sentence, clitic hopping will take place. (The

optionality is a function of dialect and idiomatic use in synchronic Spanish, as explained in Chapters 2 and 4.)

(7.20) (D-structure fragment)



The PROCL position is generated as empty in D-structure by the rules for both tensed (7.21.b) and untensed (7.21.a) constructions.

(7.21.a)  $x(pk, PT, [-tns], Inf) \rightarrow [e], \{PT = ..[pk, e]\}.$

(7.21.b)  $x(pk, PT, [+tns], Inf) \rightarrow [e], \{PT = ..[pk, e]\}.$

The CDG rule for this transformation is again a very general one (7.22). A trace is inserted in ENCL2 in place of the moved clitic(s). (If the option of suppressing clitic-hopping is taken, this rule must be removed from the grammar.)

(7.22)

$x(pk, PT1, [+tns], Inf), skip, x(ek2, PT2, [+tns], Inf) \rightarrow$   
 $x(ek2, PT1, [+tns], Inf), skip, x(ek2trace, PT2, [+tns], trace).$

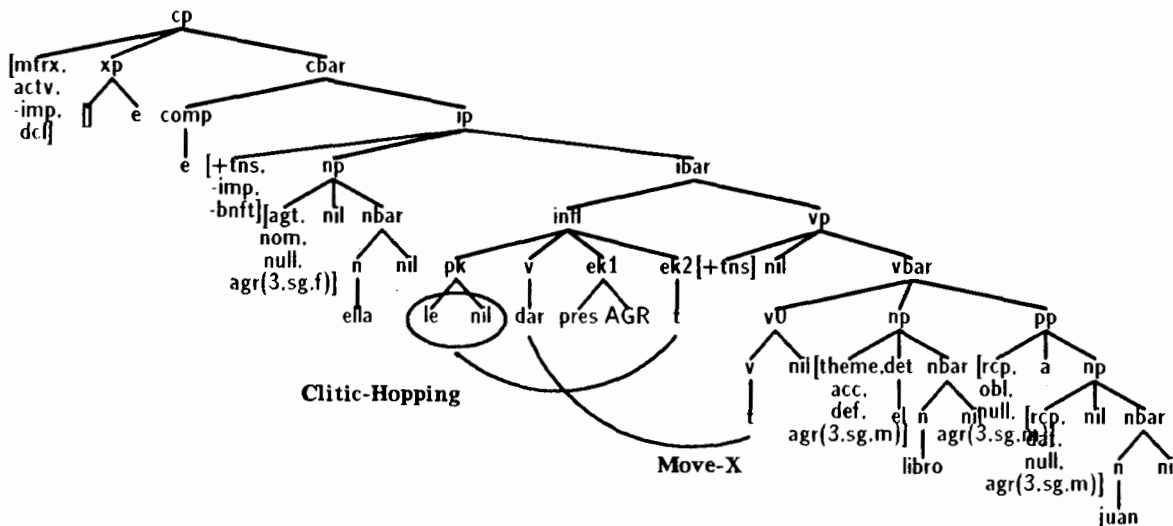
The rule (7.23) does the actual insertion of the trace.

(7.23)

$x(ek2trace, PT, [+tns], trace) \rightarrow [t], \{PT = ..[ek2, t]\}.$

The resulting S-structure for a typical sentence is given in (7.24).

(7.24) **Ella le da el libro a Juan.** (S-structure)  
*She gives the book to Juan.*



## 7.4. Clitic Promotion

Clitic Promotion is the term used to describe the apparent movement of clitics from an embedded verb to a higher verb in the syntactic structure. This phenomenon has been the subject of a growing body of literature. The discussion that follows is speculative and is included only to illustrate a possible mechanisms.

If we speculate that the movement is one of direct promotion from ENCL2 of a embedded clause to PROCL of a tensed, higher (often matrix) clause, we might propose a CDG rule (7.25) to express the transformation. (There is no evidence for cyclical movement, first to ENCL2 position, then to the PROCL position of the higher clause.)

(7.25)

$$\begin{aligned}
 &x(\text{pk}, \text{PT1}, [+tns]), \text{ skip, } x(\text{ek2}, \text{PT2}, [+tns], [[\text{K1}, \text{C1}, \text{O1}, \text{A1}], \\
 &\hspace{15em} [\text{K2}, \text{C2}, \text{O2}, \text{A2}]]) \text{ -->} \\
 &x(\text{ek2}, \text{PT1}, [+tns], [[\text{K1}, \text{C1}, \text{O1}, \text{A1}], \\
 &\hspace{15em} [\text{K2}, \text{C2}, \text{O2}, \text{A2}]]) \text{, skip, } x(\text{ek2trace}, \text{PT2}, [+tns], \text{trace}).
 \end{aligned}$$

Note the additional rule necessary if this is in fact a movement phenomenon.

(7.26)

$x(\text{ek2trace,PT,[-tns],trace}) \rightarrow [t], \{\text{PT}=\text{..[ek2,t]}\}.$

It is certainly lexically governed. However, it should be noted that there seems to be a question about the level at which clitic promotion takes place. I have no direct evidence that the mechanism is justified. It seems equally possible that Clause Union and prelexical processing (Perlmutter) are involved. If it turns out to be the case that clitic promotion is indeed a transformation, then a rule such as (7.25) might be used. If, however, clitic promotion is the result of Clause Union, then a prelexical process acting upon the semantic representation would be needed.

## 7.5. Overlapping and Nested Movement

Four different transformations have been seen above:

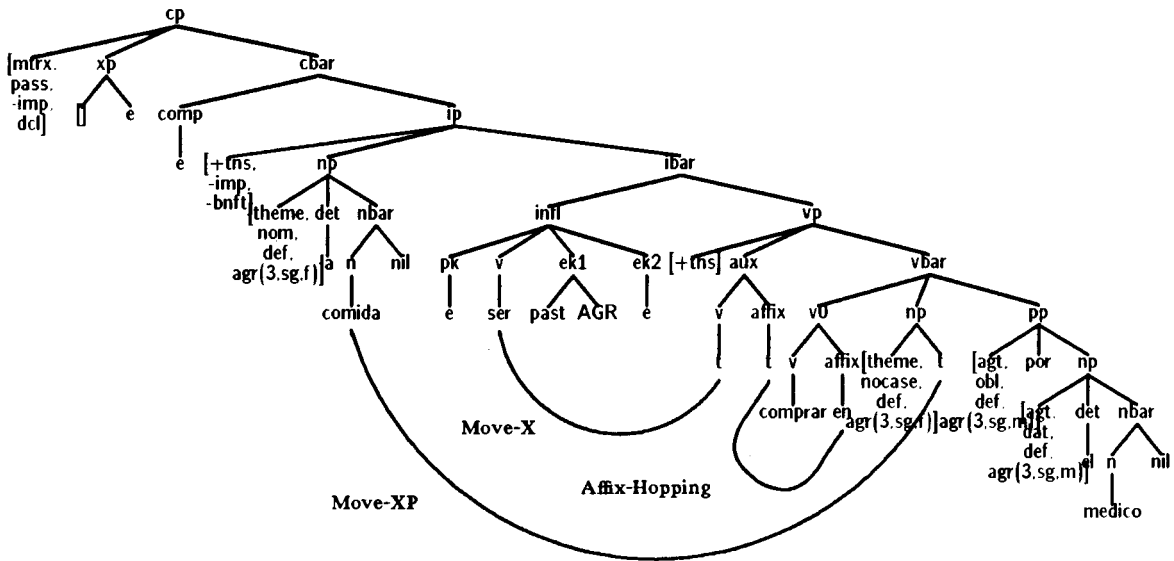
- **Move-XP:** The NP without case may move to an empty XP position.
- **Move-X:** The verb may move from Head-of-VP position to INFL.
- **Affix-hopping:** The affix *en* on the auxiliary *ser* may move to the affix-of-verb position. (A movement to the right.)
- **Clitic-Hopping:** The clitics may move from ENCL2 to PROCL position.

With CDGs it is possible to have nested transformations. Consider the S-structure (7.27) of the true passive sentence seen above. Here the Move-NP transformation has the two inflection movements nested within its scope.

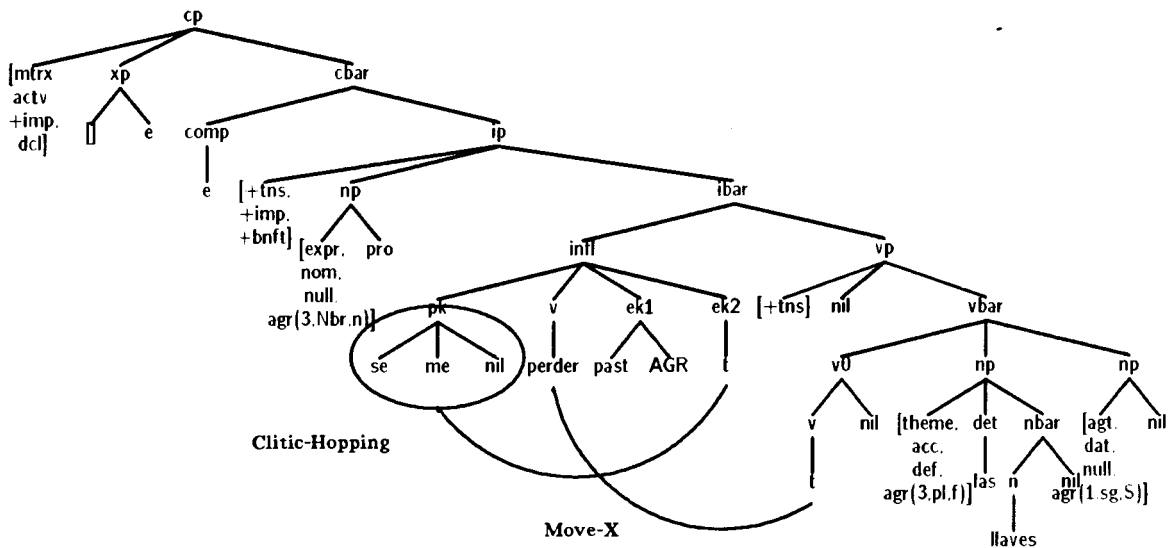
Overlapping transformations are also possible with CDGs, as they were with the original formulation of DGs. This is one of the more powerful aspects of CDGs. This cannot be done with XGs. As an example, the S-structure (7.28) demonstrates the overlapping of the Move-X transformation of raising the verb to INFL and of the clitic hopping transformation.



(7.27) **La comida fue comprada por el médico.** (S-structure)  
*The food was bought by the doctor.*



(7.28) **Se me perdieron las llaves.** (S-structure)  
*My keys got lost on me.*



## 7.6. Computing Transformations

The use of CDG rules to describe and implement linguistic transformations in the framework of GB Theory has been demonstrated in the above examples. Some comments are in order however. The transformations are rather specific. The Move-NP transformation acting as an active constraint on NPs without Case is particularly explicit. The Move-X transformation moving a verb or auxiliary to INFL is explicit about the category being moved (a verb). Similarly, the minor movement rules are distinct and explicit about the inflections which may be moved.

One of the justifications of GB Theory is explanatory adequacy. Such individual and specific transformations as those above do not have the freedom of completely free generation of sentences. This is a result of a decision to realise some computational efficiency through the use of CDG rules as active constraints.

## Chapter 8

### Remarks and Conclusions

In this dissertation I have presented a synthesis of DGs and GB theory. I have incorporated the transformational component of GB theory in a DG grammar for the generation of tensed Spanish sentences with object clitics. The new interpretation of DGs developed in [Dahl Brown et al 86] has been justified and I have developed a compiler to map DG rules under the new interpretation into Prolog Clauses.

For the logic grammar using the new interpretation I introduced an inflectional analysis of Spanish clitics. Clitic chains with the OVERT feature for each element of the chain were used to provide a general mechanism of generating pronominalised, dialectal clitic doubling, and emphatic constructions. Pre-lexical processing was postulated for the dative of interest, ethical dative, and benefactive dative constructions.

The logic grammar I have presented is able to express transformations and phrase structure rules. Move XP, Move-X, and minor movement rules were presented. Thus, the logic grammar displayed an expressive power similar to that of GB theory.

In addition, I have shown how active constraints can be incorporated in the transformational rules. In this respect, the grammar differs from GB theory which uses passive filters to rule transformed constructions ungrammatical.

I have introduced the idea of an EOS and shown how the grammar can use this set of

data structures in a generalised level of pre-lexical processing to provide an interface between a semantic representation and the grammar of generation

This chapter summarises the the contributions of the research. I shall also note some areas for future research that came to light while developing the theories and grammar presented in this dissertation.

## 8.1. Some Remarks on Automatic NL Generation

One area in which this dissertation (and all practical automatic NL generation ) differs from GB theory is the question of semantic representation. It is a hypothesis of Chomsky that syntax is autonomous. Adherents to GB theory also assume 'free generation'. Free generation can be described as allowing the generation of every possibility. The assumption is made that general principles of the grammar will rule (most) sentences ungrammatical at a later stage. Only the grammatical ones will remain. The tendency is to ignore the question of **what is to be said**. It is not conceivable that a human would generate all possible sentences of a language (grammatical and ungrammatical) until the one corresponding to the meaning sought is generated. One is led to the conclusion that the linguistic facility of the human mind has at least the ability to form a proposition (or thought) which is to be uttered. It is that proposition, in particular, that the the linguistic facility tries to place into words through the grammar of the language being spoken. We in computational linguistics have the need, for reasons of computational efficiency, to assume some internal representation of the proposition to be expressed linguistically. We call this a semantic representation. Linguists working in the framework of GB theory do not need to assume, or have not assumed, such a representation.

It seems to me that this free generation approach can lead to unreasonable hypotheses

about NL generation. Consider for example the suggestion by Rivas (discussed in Chapter 1) that all possible combinations of clitic/NP pairs would be Base generated and that some subsequent rule might delete either the clitic or the NP. In my opinion, such a suggestion lacks generality and requires specific treatment of clitic deletion. However, such an approach is perfectly reasonable within the free generation paradigm.

I have suggested another approach in Chapter 2: clitic chains with an OVERT feature. I had sought some mechanism that could make use of a semantic representation for automatic NL generation of Spanish sentences with object clitics. I had to take into account the various dialectal, emphatic, and pronominal constructions. My approach works in the grammar presented. I want to suggest that it is more general than theories such as Rivas. I am led to the conclusion that it is worthwhile in linguistic research to consider the ramifications of a semantic representation to the free generation hypothesis. Perhaps free generation is not so free after all. We are, both in Artificial Intelligence and in Linguistics, trying to model the behaviour of the mind. Surely then we must consider the notion of semantic representations (i.e. what to say: mental representations, language of thought) as a part of that model.

## 8.2. Conclusions and Contributions

### 8.2.1. Interpretation of DGs

The new formulation of DGs developed with Dahl, Boyer, and Pattabhiraman [Dahl, Brown et al 86] was explained in Chapter 6 as one requiring the simultaneous application of rules with shared variables to a derivational structure representing the abstract internal linguistic structure of a sentence.

A rule (8.1) was interpreted as the simultaneous application of the rules (8.2)

(8.1)     a, skip, b --> a', skip,b'.

(8.2.a)     a --> a'.

(8.2.b)     b --> b'.

The linguistic motivations for the new formulation were both particular and general. The particular reason was that the old formulation of DGs did not allow access to a derivational tree only to a derivational graph. In GB theory, the grammaticality of transformations depends critically upon the syntactic structure, a tree structure that may be represented by the derivational tree structure of the DG (in the new formulation but not in the old). The general reason was to provide the expressive power of the transformational component of GB theory in DGs.

Another (computational) motivation lay in the fact that DGs as originally formulated could not generate strings without going into a recursive loop. Thus, a means had to be found to interpret DG rules to give a valid linguistic interpretation and to avoid the recursive looping problem.

The generation grammar for Spanish I have presented demonstrates the linguistic validity of the new DG formulation we have developed.

## 8.2.2. Implementation of a CDG Compiler

The implementation of the compiler to map DG rules into Prolog clauses has proved the validity of the new formulation. In implementing this compiler, I have made use of the ideas of [Dahl and Abramson 84] and [Dahl 84] in their original work on DGs and have incorporated ideas for derivational context from [Dahl and St-Dizier 86] in their work on Constrained DGs under the old formulation. The original idea of a list of goals in [Dahl 84] has been changed to Saint-Dizier's idea for storing a list of rules to be applied.

The compiler allows the grammar writer to ignore the need to incorporate derivational context lists of stored rules, and DCG arguments in the grammar. These are added automatically by the compiler.

A transformational rule such as (7.10), repeated here, is automatically translated into two rules (8.3) by the compiler<sup>98</sup>

(7.10)  $x(\text{infl}, PT1, G, \_), \text{skip}, x(\text{vp}, PT2, G, W) \text{ -->}$   
 $x(\text{vp}, PT1, G, W), \text{skip}, x(\text{trace}, PT2, G, W).$

<sup>98</sup> I have not translated the unbound variable names of (8.3) into a 'readable' form, but have left them exactly as the `c prolog` interpreter displays them. Referring to the variable names of (7.10), they translate as follows:

- `_4` is `PT1`, an internal parse-tree representation.
- `_5` is `G`, the conceptual graph.
- `_6` is `_`, the 'don't care' variable.
- `_13` is `PT2`, an internal parse tree representation.
- `_14` is `W`, representing the lexical item or word.
- `_67` `x(_70)` and `_70` are derivational contexts.  
added automatically
- `_450`, `x(_453)`, and `_450` are derivational contexts.  
added automatically
- `_130`, `_59` and `_57` are stored lists of rules.  
added automatically.
- `_63`, `_514` and `_61` are stored lists of rules.  
added automatically.
- `_255`, `_256`, `_638`, `_639` are DCG list arguments.  
added automatically.

## (8.3.a)

```
x(x(_70),_67,infl,_4,_5,_6,_130,_59,_255,_256) :-
  x(_70,_67,vp,_4,_5,_14,_130,_57,_255,_256),
  conc(_57,[store(x(vp,_13,_5,_14),x(vtrace,_13,_5,_14))],_59).
```

## (8.3.b)

```
x(x(_453),_450,vp,_13,_5,_14,_63,_514,_638,_639) :-
  remove(store(x(vp,_13,_5,_14),x(vtrace,_13,_5,_14)),_63,_61),
  x(_453,_450,vtrace,_13,_5,_14,_61,_514,_638,_639).
```

### 8.2.3. The Environment of Syntax

I developed in Chapter 3 a theory of the syntax-semantics interface for a generation grammar. Procedures (called `consultcontext(XP, . . .)`) consult the EOS to find or deduce the required information for the XP being generated. The assumption here is that transformations of GB theory are meaning-preserving. Thus, the semantic representation and other components of the EOS may be consulted only in the generation of D-structure by the Base component of the grammar.

The EOS consists of:

- Lexicon
- A categorial type
- Semantic representation
- Semantic-type hierarchy
- A stack of R-expressions

This requires a set of procedures for each XP type. In developing these procedures, I included the pre-lexical processing developed for specific Spanish constructions. These procedures are definitely language-specific, as is the lexicon. Lexical entries of the EOS carry  $\Theta$ -role, Case, subcategorisation, theta-type, and animate/inanimate information for the Spanish lexicon. Distinct procedures are required for each one of the XP types {NP,PP,VP,CP,IP}



I suggested that the procedures for consulting the EOS make the grammar independent of the semantic representation used. The procedures must be able to find or deduce the information needed for a specific XP type. Thus the procedures are not independent of the semantic representation. In my generation grammar, the procedures called **consultcontext(XP, . . .)** are specific to the EOS and the semantic representation used.

#### 8.2.4. A Working Transformational DG

The NL generation grammar I have developed is the first sizable DG. The grammatical coverage is quite broad; encompassing the following types of transformations:

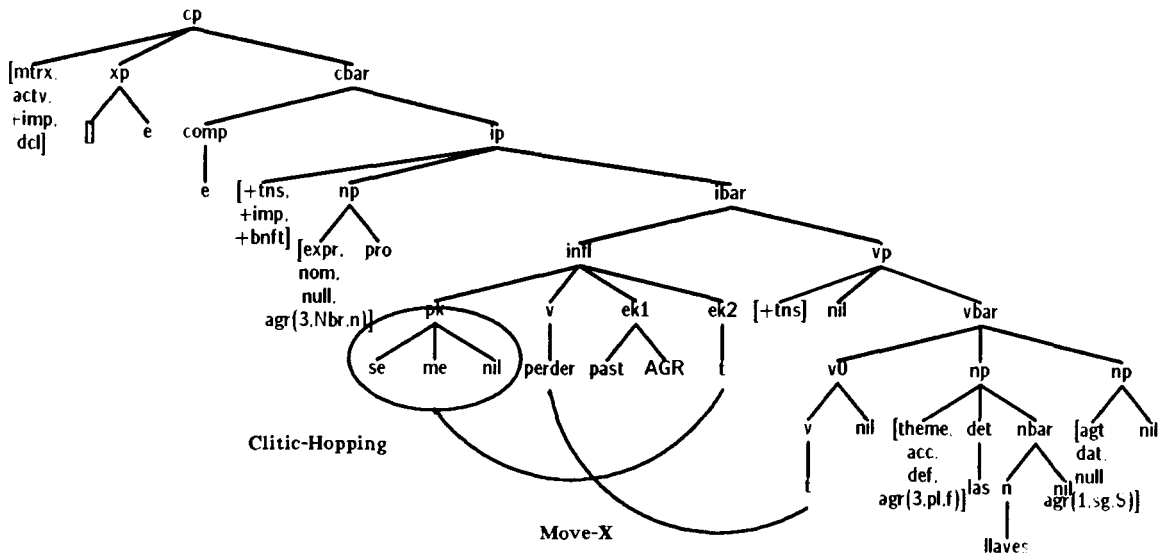
- Move-XP
- Move-X.
- Clitic Hopping.
- Affix-Hopping.

In Chapter 7 I summarised the transformational rules developed for the generation grammar and gave examples of their application to the generation of Spanish sentences with object clitics.

Sentences such as (7.27), repeated here, can be generated.

The rules for generating an XP included a list of syntactic features of the XP. Use was made of these features to ensure that the principles of Case Theory and Theta-theory are obeyed. This is was done through the use of dynamic constraints on transformations which will be discussed next.

(7.27) **Se me perdieron las llaves.** (S-structure)  
*My keys got lost on me.*



### 8.2.5. Dynamic Constraints on NL DGs

I have included dynamic constraints [Brown, Pattabhiraman et al 86] on some transformations. Transformations in DG rules may be quite general. For example, the rule for clitic-hopping (7.22), repeated here for convenience, was very general and applied to all tensed sentences.

(7.22)

$x(pk, PT1, [+tns], Inf), skip, x(ek2, PT2, [+tns], Inf) \rightarrow$   
 $x(ek2, PT1, [+tns], Inf), skip, x(ek2trace, PT2, [+tns], trace).$

On the other hand, the rule (7.4), repeated below, embodied a dynamic constraint on the transformation. The dynamic constraint in this case corresponded to the Case Filter of GB theory. A filter that applies after the transformational component of GB theory. The dynamic constraint in (7.4) prevents the transformation from occurring.

```

(7.4)
xp(np,PT1,G,[no-tr,nom,Def1,Agr1,Ovrt1,Animate1]),
  skip,
    xp(np,PT2,G,[theme,nocase,Def2,Agr2,Ovrt2,Animate2])
-->
xp(np,PT1,G,[theme,nom,Def2,Agr2,Ovrt2,Animate2]),
  skip,
    xp(nptrace,PT2,G,[theme,trace,Def2,Agr2,Ovrt2,Animate2]).

```

The need to constrain some transformations dynamically is due solely to the computational consideration of efficiency (The choice is to allow the transformation to apply universally and then to rule some of the transformations ungrammatical at a later stage of generation or to embody some of the filters of GB theory in the transformational rule. In doing so, the transformation is prevented from occurring in the first place.)

### 8.2.6. Generation of Passives and Actives

The generation grammar I have presented makes use of the non-deterministic proof procedure of Prolog. Both active and passive versions of a sentence are generated automatically (unless one is specifically specified in the original call to the generation grammar).

- (8.4.a) El médico compró la comida.  
*The doctor bought the food.*
- (8.4.b) La comida fue comprada por el médico.  
*The food was bought by the doctor.*

To the best of my knowledge, no previous logic grammar has succeeded in generating both actives and passives from the same semantic representation. This may be true of all previous work in automatic NL generation.

I have used the general mechanism of pre-lexical processing to effectively re-arrange the argument structure of the proposition in the semantic representation. This mechanism,

together with specific lexical entries for verbs which allow passivisation, permitted the non-deterministic generation of both voices

### 8.2.7. Non-Deterministic Generation

The nondeterminism of the generation was used in more than one way. Pronominalised and non-pronominalised NPs were both generated through automatic backtracking, using the OVERT feature. The example given in Chapter 5, repeated here, shows this clearly.

- (5.31.a) **El médico la compró.**  
*The doctor bought it.*
- (5.31.b) **El médico compró la comida.**  
*The doctor bought the food.*
- (5.31.c) **La compró.**  
*He bought it.*
- (5.31.d) **Compró la comida.**  
*He bought the food.*
- (5.31.e) **La comida fue comprada por el médico.**  
*The food was bought by the doctor.*
- (5.31.f) **La comida fue comprada.**  
*The food was bought.*

### 8.2.8. Base Generation of Clitics

I have used three working hypotheses in Chapter 2. Firstly, that a clitic chain is Base generated in Spanish for each and every A-position. Secondly, that both the clitic and the A-position have a distinct OVERT feature. Thirdly, that at least one element of a clitic chain in a tensed construction must have its OVERT feature set to [+ovrt].

I used these working hypotheses in the development of the generation grammar to demonstrate that they were adequate for the generation of pronominalised NPs, clitic doubling constructions, emphatic constructions, and dialectal doubling constructions.

The generation grammar allows the generation of all of these constructions. Thus a sentence, if not constrained, may be generated in several versions.

- (8.5.a) **Juan lee el libro.** (Real Academy Spanish)  
*Juan reads the book.*
- (8.5.b) **Juan lo lee el libro.** (clitic doubling, dialectal/emphatic)  
*Juan reads the book.*
- (8.5.c) **Juan lo lee.** (object pronominalisation)  
*Juan reads the book.*
- (8.5.d) **Lee el libro.** (null-subject pronominalisation)  
*Juan reads the book.*
- (8.5.e) **Lo lee.** (object and null-subject pronominalisation)  
*Juan reads the book.*

My analysis of Base generation of clitics differs from other reported research. In Chapters 1 and 2 I pointed out that differing approaches to the status of clitics with respect to the Binding Condition and the Empty Category Principle have been proposed. All of these required separate analyses of the constructions mentioned above. I have proposed the clitic chain analysis for reasons of computational simplicity. It seems to me that an argument can be made for the my analysis on the grounds of explanatory adequacy. One mechanism of Base generation that replaces three or four distinct mechanisms has much greater generality, and hence explanatory adequacy. I require the OVERT feature for the elements of the Base generated clitic chains. (The inflectional analysis of clitics is not central to the clitic chain analysis, so long as clitics are found to have no status themselves with respect to the Binding Theory.)

### 8.2.9. Analysis of Clitics as Inflections

In Chapter 4, I hypothesised that the Base generation of clitics may be treated with an inflectional analysis: clitics as verbal inflections. I suggested that ENCL1, ENCL2 and PROCL positions for inflections of Spanish verbs exist. In worked shared with Sempere, we argued that there are compelling reasons to treat clitics as inflections, and not as a separate categorial type (and in particular not as pronouns). We based our arguments on the evidence that clitics are not pronouns, on evidence of the proximity to the verb of

clitics on dialectal evidence that indicates clitics may occur in tensed constructions in enclitic position, and on arguments of explanatory adequacy.

We noted that other researchers were divided in their opinions as to how to treat the Base generation of clitics. Our idea that object clitics are base generated in INFL in Spanish runs parallel to Roberge's theory that colloquial French subject clitics are Base generated in INFL

We suggested that clitic-hopping from ENCL2 to PROCL position does not leave a trace because inflections have no status with respect to the Binding Condition.

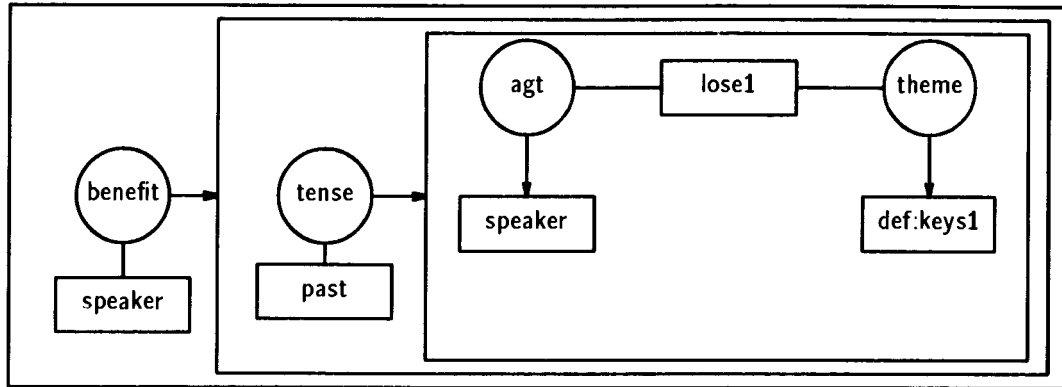
I suggested that the  $X^0$  verb movement from VP to INFL suggested in Barriers [Chomsky 86] could be used in this connection without disturbing the inflectional analysis. This movement was implemented with a CDG transformation in the generation grammar for Spanish

#### **8.2.10. Representation of Special Clitic Constructions**

Some of the clitic constructions to be generated required semantic representations that go beyond the normal predication plus arguments structure of a straightforward utterance. In Chapter 2, I suggested these constructions (benefactive datives, datives of interest, and ethical datives) share a common form of semantic representation; namely a higher relation (benefit, interest, and affront respectively) one of whose arguments is the sentential proposition

In addition, I suggested that all three require the addition of an extra argument at the syntactic level to express the meaning correctly. In Chapter 5 I showed how such an extra argument could be added in the prelexical processing of the VP

- (2.31) **Se me perdieron las llaves.**  
*The keys got lost on me.*



### 8.2.11. Pre-Lexical Processing

I suggested in Chapter 2 that a level of pre-lexical processing is necessary for the generation of several constructions. These included dative constructions such as the benefactive dative, ethical dative, and dative of interest, and true passives.

For these dative constructions, I showed that they could be thought of as sharing a common form of semantic representation. I proposed that they have a semantic representation with a higher relation (**benefit**, **affront**, or **interest**) involving one of the arguments of the proposition of the sentence.

In the generation grammar, I implemented pre-lexical processing within the processes that consulted the EOS. The success of these processes in generating benefactive datives and true passives demonstrates the feasibility of this level of processing, and the corresponding usefulness of the semantic representations.

True passive constructions were generated with corresponding active sentences using the

non-deterministic proof procedure of Prolog. I assumed that verbs that passivise have a separate lexical entry for the passive form, but that both active and passive sentences were generated from the same semantic representation.

### 8.3. Remarks on Some Open Questions

There were some questions examined but left only partially answered in the earlier Chapters. I want to summarise them here and suggest that they are worthy of further examination.

#### 8.3.1. The Theta Criterion and Chain Composition

In Chapter 2 I raised the question of whether or not the *recipient* relation actually exists in sentences such as (2.22), repeated here for convenience. There is no question about the grammaticality of these possessive dative constructions. Nor are they unique to Spanish.

- (2.22) **Los amos no le pagaron el sueldo.**  
*The bosses did not pay him his wages.*

The difficulty raised was that if the *recipient* relation actually exists in the semantic representation, then the dative clitic [le] appears to have two  $\Theta$ -roles through co-indexation and chaining. The Theta Criterion of GB theory permits the assignment of only one  $\Theta$ -role to an A-position. I suggest that this question be examined very closely with respect to assignment of  $\Theta$ -roles to  $\bar{A}$ -positions. In this case, the definite article [el] is co-referential to an  $\bar{A}$ -position and appears to be assigned the *possessive*  $\Theta$ -role which is also co-referential to the clitic [le].

I have assumed that the two indexations (clitic chain and possessive co-reference) are not composed at LF. More work could be done in this area. The fundamental question



is whether the Theta Criterion should allow chain composition of this sort. If not, I suggest that the Theta Criterion must be re-examined with respect to its applicability to  $\Theta$ -roles assigned to  $\bar{A}$ -positions or assigned by prepositions (such as the possessive preposition *de* in (2.23.b)) to their complements

(2.23.a) **Los médicos le cerraron los ojos.**  
*The doctors closed his eyes.*

(2.23.b) **Los médicos le cerraron los ojos de Juan.**  
*The doctors closed Juan's eyes.*

### 8.3.2. Temporal Relations and Mood, Tense, Voice, and Aspect

One area of concern for automatic system for NL generation, especially one which might form part of a NL translation system, must be that there is a great variation in the use of Tense, Voice, Aspect, Mood between languages. This is especially true in the use of compound tenses in some languages (Spanish, for example.)

I did not attempt to incorporate Mood and Aspect into my generation grammar for Spanish (Nor was this the intent of the research.) However, it has become evident during the research, that the interpretation of semantic representation must eventually involve these areas. There are really two areas here. One is the inclusion in semantic representations for utterances of temporal relations between time-of-utterance, time-of-occurrence and time-of-reference. The other is the interpretation of semantic representations with these temporal references by a generation grammar.

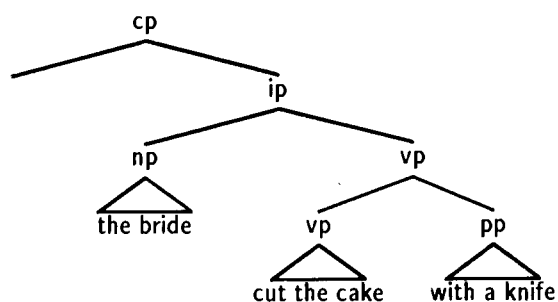
I feel that it is time that computational linguists turn their attention to the inclusion of temporal references in semantic representations for sentences (and not just frames or scripts). The application of grammars such as mine to practical systems of NL translation requires work in this area to be really useful.

### 8.3.3. Expression of Adjunction in Logic Grammar Rules

There are two kinds of adjunction in GB theory. One is the adjunction of  $\bar{A}$ -position XPs in the Base generation of sentences. The other is adjunction transformations.

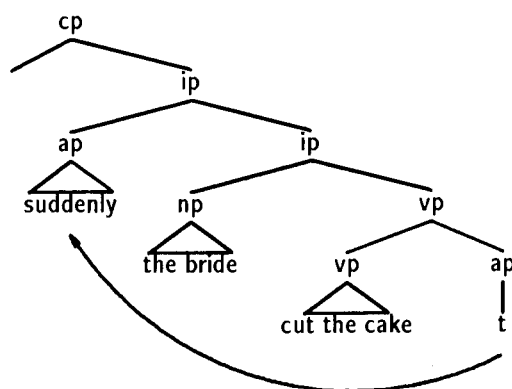
Adjunction in Base generation is illustrated in (8.6) where the AP **with a knife** is adjoined to the VP in the sentence. Question of adjoined modifiers. Adverbial and Adjectival phrases.

(8.6) **The bride cut the cake with a knife.**



Adjunction movement is illustrated in (8.7) where the AP has undergone a movement transformation to adjoin to IP. Question of adjunction movement.

(8.7) **Suddenly the bride cut the cake.**



The question I want to pose is, how can these two types of adjunction be expressed in logic grammar rules? Should they be expressed in different ways? Can we, for example, separate the adjunction (which is common to both) from the movement transformation?

In my generation grammar, there is no provision for base generation of adjoined XPs. If such a mechanism were to be included, what would be the role of the `consultcontext(AP, . . .)`, for example, if an AP were to be adjoined to a VP, and how would this process interact with the `consultcontext(VP, . . .)` process? Would any interaction at all be necessary? desirable?

#### 8.3.4. Clitic Promotion

I pointed out in Chapter 7 that there is a question whether clitic promotion is a transformation or a pre-lexical process. There is a growing body of literature in linguistics on this question. Clitic promotion occurs in several languages. No principled mechanisms have been clearly established. The Clause Union work of Perlmutter may lead to a principled explanation of this phenomenon. The Spanish data is very confusing [Brown 86]. It is certain that GB theory at present provides no mechanism for clitic promotion other than  $\text{Move-}\alpha$  transformations. However, no constraint mechanism on these transformations is clearly established. In the meantime, I suggest the prelexical processing approach is the most promising for computational applications.

#### 8.3.5. Projection

The Projection Principle of GB theory states that features projected from the lexicon are preserved at all levels. I have used projected features in my generation grammar for Spanish. Projection is actually accomplished in the top-down proof procedure of Prolog through automatic backtracking. If features do not unify, then backtracking tries to find another rule in which features will unify.

The grammar is based on phrase structure rules as the basic method of stating the syntactic structure of the language. Projected features are taken to be a secondary

consideration to the phrase structure. They act as constraints on the phrase structure. I wonder whether the roles can be reversed? That is can a logic grammar be based on projection rules, with phrase structure as a constraint?

One of the advantages of such a grammar might be computational efficiency. Less backtracking might be required. What I now see as excessive, but unavoidable, backtracking in a phrase structure based grammar might be avoided.

#### **8.4. Final Comments**

This dissertation has demonstrated the potential of the synthesis of DGs and GB theory in NL language processing. I believe that the way is now clear to applying the principles developed to NL analysis, NL translation, and QA systems. I have pointed out the areas that touch on this work which, in my opinion, merit further research.

## Appendix A

### A Glossary of Some Concepts of GB Theory

**The Base component of Universal Grammar:** The Base component of Universal Grammar consists of the categorial component and the lexicon. The rules of the categorial component obey some variety of X-bar theory. The lexicon specifies the morphological and phonological structure of lexical items, the categorial features of the item, its contextual features and its subcategorisation.

The Base component generates D-structure.

**Binding Theory:** The co-reference of NPs is represented in GB theory by co-indexation. An NP is bound if it is co-indexed with a governing argument. An NP which is not bound is free. The principles proposed in LGB, [Chomsky 82a] are:

- A A anaphor must be bound in its governing category.
- B A pronominal must be free in its governing category.
- C A referential expression (R-expression) is free.

**Bounding theory (Subjacency):** Bounding theory serves to constrain the movement of constituents out of XPs. The theory is language specific in its boundaries to movement. Bounding theory appears to provide some flexibility in the severity of violations. This aspect of GB theory is currently under scrutiny. In *Barriers*, [Chomsky 86], an attempt is made to subsume Bounding theory under the notion of Government.

**Case Theory:** Case Theory in GB states simply that: *Every phonetically realised NP*

*must have Case assigned.* Case may be assigned in either of two ways. Inherent Case marking applies in the Base component. Structural Case marking applies at S-structure. In addition, Case filters operate to rule ungrammatical some constructions.

As we have seen, Case is assigned to A-positions in D-structure by projection from the Lexicon as the EOS is consulted from VP. Put simply, Case is assigned to an NP by a category (e.g. a verb or preposition) that governs it. The following Case-marking rules for English are cited in [Chomsky 82a]:

- An NP is assigned objective Case if governed by a transitive verb or by a preposition. This is a bit of a circular definition, as Chomsky defines a transitive verb as one which assigns Case to its complements.
- An NP is assigned nominative Case if governed by +TNS (tense).
- An NP is assigned genitive Case if governed by the feature POSS.

Several filters exist which rule some constructions ungrammatical at S-structure. The NP trace filter says that the trace of a NP movement cannot be Case marked. The PRO filter says that PRO cannot be governed. The Case filter says that no overt NP can be without Case. The Case Conflict Filter says that no NP can have more than one Case assigned to it. These are discussed in LGB [Chomsky 82a].

**Control theory:** The choice of antecedent of the empty category PRO is determined by Control theory.

**Empty Categories:** The question of the nature of empty categories in Government and Binding theory is the subject of a great deal of research. In addition to [Chomsky 82a], the discussion of empty categories in [Bouchard 83] is very useful.

Currently the following empty categories are generally recognised:

- **trace:** A trace always has an antecedent, namely the moved category whose position the trace occupies. A trace must be governed.

- **PRO:** The empty category PRO occurs in clausal complements. PRO need not have an antecedent, and when an antecedent does occur it has a  $\theta$ -role independent of that of PRO. PRO is not governed.
- **pro:** The empty category pro is the non-overt subject of null-subject languages such as Spanish and Italian.

**The Empty Category Principle:** The Empty Category Principle (ECP) states that every trace must be properly governed. A good discussion of the ECP is to be found in [Chomsky 82b]. Also see [Bouchard 83].

**Explanatory Adequacy:** Chomsky has proposed three criteria for explanatory adequacy for a grammar of a language (paraphrased here from [Radford 81]):

- The grammar correctly predicts which sentences are and are not well formed in the language
- The grammar correctly describes the structure of the sentences of the language.
- The grammar does so in terms of a highly restricted set of optimally simple, universal, and maximally general principles of mental computation learnable by a child

Radford has a useful discussion of this, as does *Lectures on Government and Binding* [Chomsky 82a]. The concern is that a grammar ought to be able to explain language acquisition by a child in a few years with a poverty of stimulus.

**Government theory:** Proper government may occur two ways; either through lexical government or through antecedent government. In *Barriers* [Chomsky 86] the principles of government are re-examined and criteria for proper government are developed. These criteria are more intricate and more universal than the earlier definition of government based on c-command.

The c-command definition of government given in [Chomsky 82a] states: In the structure  $[\gamma \beta \quad \alpha \quad \beta]$ ,  $\alpha$  governs  $\beta$  iff  $\alpha$  is an immediate constituent of  $\gamma$  and, where  $\phi$  is a maximal projection, if  $\phi$  dominates  $\beta$  then  $\phi$  dominates  $\alpha$ .

**Maximal Projection:** A category XP is the maximal projection of a lexical category X. Thus the NP immediately dominating a noun, or trace, is the maximal projection of that noun. A N-bar immediately dominating a noun is a projection of that noun, but not a maximal projection.

**The Projection Principle:** The Projection Principle states that *Representations at every syntactic level are projected from the Lexicon*. That is, the subcategorisation properties and thematic relations of lexical items must be observed at every level.

**The Theta Criterion:** The Theta Criterion as stated in [Chomsky 82a] states that:

- Each argument bears one and only one  $\Theta$ -role.
- Each  $\Theta$ -role is assigned to one and only one argument.

**Theta Theory** Theta Theory is concerned with the thematic roles associated with grammatical functions (subject, object, etc. of a verb). These positions are normally called *A-positions*. Chomsky believes that all and only those A-positions that are assigned  $\Theta$ -roles are lexically filled at D-structure (assuming that the phonetically null element PRO may bear a  $\Theta$ -role and be considered a 'lexical' item).

**The Transformational Component of UG:** Move- $\alpha$  is the principal mechanism of the Transformational component. There are two types of movement allowed.

- **XP-movement:** A maximal projection (XP) may move only to an empty XP position or to adjoin to an XP position.
- **X<sup>0</sup>-movement:** A head (X) may move only to an empty head position.

In addition, minor movement rules of Affix-hopping, Have-Be raising, (and Do-support in English and Swiss-German) provide an explanation of verb and auxiliary morphology. At this time, since the publication of *Barriers* [Chomsky 86], the status of affix-hopping is in doubt.



The Transformational Component generates S-structure through transformations on D-structure.

**Universal Grammar (UG):** The theory of Universal Grammar according to [Chomsky 82a] must be compatible with the diversity of existing grammars of Natural Language. It must also be constrained enough to account for the development of these grammars in the human mind with only a very limited input of data, utterances from other humans.

On the basis of these two criteria, Chomsky expects the theory of UG to be highly structured and based on a small number of principles that severely constrain the set of possible grammars. He also expects a set of parameters, set on the basis of input utterances that determines the form of individual languages.

**X-bar Theory:** There is a great deal of evidence that syntactic constituents exist. X-bar syntax recognises the existence of a hierarchy of syntactic constituents. Several schema have been proposed. See [Radford 81] for a discussion of schema which introduce up to six bar-levels: X, X-bar, X-double-bar, . . . , X-with-n-bars. The analysis of [Stowell 81] provides a more elegant treatment of X-bar theory with just the three categorial types: X, X-bar, and XP.

## Appendix B

### A Compiler for Constrained Discontinuous Grammar Rules

```
/* -----  
/* synapse3      Compiler for  CDG Rules using goal-list ideas.  
/*  
/*              author: Charles  
/*              created: 1987 February 09  
/*              modified: 1987 February 18  
/*              modified: 1987 February 19 (automatic context)  
/*              modified: 1987 February 20 (elegance)  
/*              modified: 1987 February 21 (makePT)  
/* NOTES  
/* Derived from the synapse (Dec 1985) and synapse2 (Jan 1986) compilers.  
/*  
/* More that one goal before and after the skip in the body allowed..  
/*  
/* This implementation requires that ALL phrase-structure-rule files be  
/* compiled using the 'compile' predicate below.  
/*  
/* The arguments for the stored goal-list are added automatically.  
/*  
/* The arguments for the derivational context are added automatically  
/* -----
```

```

/* ----- normal dcg clauses ----- */
synapse3( (A-->B). Clause ) :- !,
    newterm(B.VL,NVL,Context,L0,L1,NewB),
    makePT(A,NVL,PT),
    newhead(A,PT,Context,L0,L1,NewA). !,
    expand_term( (NewA-->NewB). Clause).

/* ----- constrained discontinuous grammar clause ---- */
synapse3( (A.skip,B-->C). Clause1. Clause2 ) :- !,
    findSkip(C,A1,B1),
    newterm(A1.VL,NVL,Context,L0,L1,NewA1),
    makePT(A,NVL,PT),
    newhead(A,PT,Context,L0,L2,NewA). !,
    (C1=(NewA-->NewA1,{conc(L1,[store(B,B1)],L2)})),
    expand_term(C1.Clause1).

    newterm(B1.VLB,NVLB,ContextB,L4,L5,NewB1),
    makePT(B,NVLB,PTB),
    newhead(B,PTB,ContextB,L3,L5,NewB). !,
    (C2=(NewB-->{remove(store(B,B1),L3,L4}).NewB1)),
    expand_term(C2.Clause2).

```

```
/* findSkip(C,A1,B1): In clause body C the predicates (A1) preceding
/* the skip and the predicates (B1) following the skip are identified. */
```

```
findSkip( (A1.skip,B1), A1,B1) :- !.
findSkip( (F.S,R), (F,AR), B1) :- findSkip( (S,R), AR, B1).
```

```
/* newterm(Term,VarList,NewVarList,Context,Le,Ls,NewTerm):
/* A parse tree variable (PT) and derivational context (Context)
/* variable are automatically added to each predicate (Term) of
/* the RHS of a clause. (recursively).
/* In addition, the I/O lists (Le and Ls) of stored goals are
/* added to the variable lists of each predicate.
/* VList is the input list of PT variables collected so far on
/* the RHS, NewVList is the output list of same.
```

```
newterm(Term,VList,VList,Context,Ls,Ls,Term) :- Term='!', !.
newterm(Term,VList,VList,Context,Ls,Ls,Term) :- Term=[], !.
newterm(Term,VList,VList,Context,Ls,Ls,Term) :- Term=..['|R], !.
newterm(Term,VList,VList,Context,Ls,Ls,Term) :- Term=..'='R], !.
newterm({Term},VList,VList,Context,Ls,Ls,{Term}) :- !.
newterm( (First,Rest), VList,NewVList,Context,Le,Ls, (NewFirst,NewRest)) :-
    newterm(First,VList,TempVList,Context,Le,L1,NewFirst),
    newterm(Rest,TempVList,NewVList,Context,L1,Ls,NewRest).
newterm(Term,VarList,NewVarList,Context,Le,Ls,NewTerm) :-
    Term=..[Head|Args],
    conc([PT,Context],Args,TempArgs),
    addArgs(TempArgs,Le,Ls,NewArgs),
    conc(VarList,[PT],NewVarList),
    NewTerm=..[Head|NewArgs].
```

```

/* makePT(Term,VarList,PT):  the PT is constructed from the variable
/*                          list (VarList) associated with the functor
/*                          heads of the RHS of the Horn Clause and the
/*                          functor name (Head) of the LHS (Term).
makePT(Term,VarList,PT) :- Term=..[Head|Args], PT=..[Head|VarList], !.
makePT(Term,VarList,PT) :- Term=..[Head|Args], PT=..[Head|[VarList]].

```

```

/* newhead(Term,PT,Context,Le,Ls,NewTerm): the PT, the Context, and the
/*                                          I/O lists are added to a(x.Y.Z) which becomes:
/*                                          head(PT,Context,x,Y,Z,Le,Ls)
newhead(Term,PT,Context,Le,Ls,NewTerm) :-
    Term=..[Head|Args],
    conc([PT,Context],Args,TempArgs),
    addArgs(TempArgs,Le,Ls,NewArgs),
    NewTerm=..[Head|NewArgs].

```

```

/* addArgs(List,L0,L1,NewList):  The arguments L0 and L1 are appended to
/*                              the input list (List) to form NewList.
addArgs([],L0,L1,[L0,L1]).
addArgs([H|T],L0,L1,[H|R]) :- addArgs(T,L0,L1,R).

```

```

conc([],L,L) :- !.
conc([H|X],Y,[H|Z]) :- conc(X,Y,Z).

```

```

remove(X,[X|Y],Y) :- !.
remove(X,[X1|Y],Z) :- remove(X,Y,Z1),conc([X1],Z1,Z).

```

```
/* ----- convenience predicates - compile dynamically ----- */
```

```
compile(Infile) :- see(Infile),get_cdg_rule(Infile),seen.
```

```
get_cdg_rule(Infile) :- read(Rule), process_cdg_rule(Rule).
```

```
process_cdg_rule(Rule) :- Rule=done,!.
process_cdg_rule(Rule) :- synapse3(Rule,Clause1,Clause2),
```

```
    assertz(Clause1),
```

```
    assertz(Clause2),
```

```
    get_cdg_rule(Infile).
```

```
process_cdg_rule(Rule) :- synapse3(Rule,Clause),
```

```
    assertz(Clause),
```

```
    get_cdg_rule(Infile).
```

```
/* ----- convenience predicates - compile to file ----- */
```

```
compile(Infile,Outfile):-see(Infile),tell(Outfile),
```

```
    find_cdg_rule(Infile),
```

```
    seen,told.
```

```
find_cdg_rule(Infile) :- read(Rule), translate_cdg_rule(Rule).
```

```
translate_cdg_rule(Rule) :- Rule=done,!.
translate_cdg_rule(Rule) :- synapse3(Rule,Clause1,Clause2),
```

```
    write(Clause1),write('.'), nl, nl,
```

```
    write(Clause2),write('.'), nl, nl,
```

```
    find_cdg_rule(Infile).
```

```
translate_cdg_rule(Rule) :- synapse3(Rule,Clause),
```

```
    write(Clause),write('.'), nl, nl,
```

```
    find_cdg_rule(Infile).
```

## Appendix C

### Arguments that Clitics are not NPs

This appendix gives the arguments that have been presented elsewhere, primarily in [Strozer 76]. They should be seen as additional support for the theory presented in Chapter 4.

#### C.1. Morphological differences

The morphological forms of clitics and pronouns in Spanish are quite distinct. All three persons have differences. This is easy to see in the following table. The extended lexicalist hypothesis in [Jackendoff 72] holds that transformations do not change morphological items. Thus a pronoun form cannot turn into a clitic form and vice versa in the course of generation.

##### (C.1)

features	pronoun-form	clitic-form
1st per pl	mí	me
1st per sg masc	nosotros	nos
1st per sg fem	nosotras	nos
2nd per sg	ti	te
2nd per pl masc	vosotros	os
2nd per pl fem	vosotras	os
3rd per reflexive	sí	se
3rd per neuter	ello	lo (accusative)
3rd per neuter	ello	le (dative)
3rd per sg masc	él	lo (accusative)

3rd per sg fem	ella	la (accusative)
3rd per sg masc	él	le (dative)
3rd per sg fem	ella	le (dative)
3rd per pl masc	ellos	los (accusative)
3rd per pl fem	ellas	las (accusative)
3rd per pl masc	ellos	les (dative)
3rd per pl fem	ellas	les (dative)

## C.2. NP forms do not distinguish Case

In table (C.1), it is evident that third person non-reflexive clitics distinguish Case, while pronouns and NPs never do. This illustrates another difference between clitics and NP forms.

## C.3. NPs take stress

Pronouns and NPs in Spanish may take emphatic or contrastive stress. In (C.2) the object NPs, whether pronominal (C.2.b) or not (C.2.a), may be stressed.

- (C.2.a) **No veo más que a JUANITA.**  
*I do not see anyone but JUANITA.*
- (C.2.b) **No veo más que a ELLA.**  
*I do not see anyone but HER.*

This is impossible with clitics. The clitic *la* in (3.3) may never be stressed.

- (C.3) \***No LA veo.**  
*I do not see HER.*

## C.4. Phonological considerations

A clitic and its verb form a single word. As a single phonological unit, both the verb and the clitic must be present. It is impossible to have a clitic in the absence of a verb.

- (C.4.a) **¿A quién viste ayer?** from Strozer (3.12)  
*Who did you see yesterday?*



In answer to (C.4) one might use a noun (C.5.a) or a pronoun (C.5.b), but never a clitic alone (C.5.c).

- (C.5.a)     **a Juanita.**  
 (C.5.b)     **a ella.**  
 (C.5.c)     **\*a la.**

Additional evidence that a clitic and its verb form a single phonological unit comes from conjoined structures (C.6). In the structures with auxiliary verbs, both the auxiliary **ha** and the clitic **lo** must be present in the second clause. The absence of one or the other is unacceptable (C.6.b.c).

- (C.6.a)     **Juan lo ha comprado y lo ha leído.**  
               *Juan bought it and read it.*  
 (C.6.b)     **\*Juan lo ha comprado y lo leído.**  
 (C.6.c)     **\*Juan lo ha comprado y ha leído.**

## C.5. Surface order of NPS

The surface order of NPs may vary (C.7.a.b). The order of clitics is inviolable. Thus, (C.7.c) is grammatical but (C.7.d) is not.

- (C.7.a)     **Juan dió el libro a María.**  
               *Juan gave the book to María.*  
 (C.7.b)     **Juan dió a María el libro.**  
               *Juan gave María the book.*  
 (C.7.c)     **Juan se lo dió.**  
               *Juan gave her/him it.*  
 (C.7.d)     **\*Juan lo se/le dió.**

## C.6. NPs take modifiers

Only NPs (C.8.a.b), and not clitics (C.8.c.d), may take modifiers.

- (C.8.a) **Hablé con ellos dos.** from Strozer (3.16)  
*I spoke with those two.*
- (C.8.b) **Escribí dos libros largos.**  
*I wrote two big books.*
- (C.8.c) \***Les dos hablé.**
- (C.8.d) \***Los largos escribí.**

## C.7. NPs may be conjoined

Both pronominal and non-pronominal NPs may be conjoined (C.9.a.b). This is not the case for clitics which may never be conjoined (C.9.c.d).

- (C.9.a) **Hablé con ellos y con Juanita.**  
*I spoke with them and with Juanita.*
- (C.9.b) **Escribí dos libros largos y dos cartas.**  
*I wrote two big books and two letters.*
- (C.9.c) \***Les y le hablé.**
- (C.9.d) \***Los y las escribí.**

## C.8. Transformations apply only to NPs

Transformations such as topicalisation and passive NP raising apply to NPs. Thus, in (C.10.a.c) it is possible to topicalize the object NP to get (C.10.b.d).

- (C.10.a) **Vio Pablo a Lola.**  
*Pablo saw LOLA.*
- (C.10.b) **A Lola la vio Pablo.**  
*LOLA Pablo saw.*
- (C.10.c) **Vio Pablo a ELLA.**  
*Pablo saw HER.*
- (C.10.d) **A ELLA la vio Pablo.**  
*HER Pablo saw.*

Such transformations do not apply to clitics. The passive subject of (C.11.a) is the NP [Ella]. It is not possible to have a clitic subject such as (C.11.b).

- (C.11.a) **Ella fue vista por Pablo.**  
*She was seen by Pablo.*
- (C.11.b) \***La fue vista por Pablo.**

A similar argument may be made for the clitic promotion phenomenon. Only clitics undergo promotion (C.12.a.b). It is not possible for NPs to be promoted (C.12.c.d).

- (C.12.a) **Enrico quiere hacerlo.**  
*Enrico wants to make it.*
- (C.12.b) **Enrico lo quiere hacerlo.**  
*Enrico wants to make it.*
- (C.12.c) **Enrico quiere hacer una mesa.**  
*Enrico wants to make a table.*
- (C.12.d) **\*Enrico una mesa quiere hacer.**

### C.9. Only clitics may double NPs

Only clitics may double NPs (C.13.a). NPs may not double NPs (3.12.b).

- (C.13.a) **Mafalda lo vi a Enrico.** (Rio de la Plata dialect)  
*Mafalda saw Enrico.*
- (C.13.b) **\*Mafalda (a) él vi a Enrico.**

## Acknowledgement

T. Pattabhiraman and Verónica Dahl contributed to this document in many useful discussions and through their incisive comments. Patrick Saint-Dizier's comments and advice were very provocative and invaluable in clarifying my ideas. T. Pattabhiraman and Pierre Massicotte are responsible for the programming of the processgraph algorithm. Both Michel Boyer and John Knowles contributed insights and many long hours to discussions of the computational and linguistic problems. Dick DeArmond contributed to my thoughts on semantics especially Frits Swinkels developed the program to automatically generate the diagrams in the dissertation.

Many others contributed time, help, inspiration, and encouragement. Thanks especially to Lou Hafer, Wo-Shun Luk, Art Liestman, Binay Battacharaya, Brian Funt, Jim Delgrande, Ted Sterling, Nick Cercone, Elma Krbavac, Doreen Godwin, Ethel Inglis, Diane Calnek, and Janis Horne. I hope that those I have forgotten to mention will forgive me.

Gracias especialmente a los Aymaras, con quienes me trajo mucho placer trabajar.

## References

- [Abramson 84] Harvey Abramson.  
Definite Clause Translation Grammars.  
In *Proceedings Logic Programming Symposium, Atlantic City*, pages  
233-240. IEEE, 1984.
- [Aoun 81] Joseph Aoun.  
*The Formal Nature of Anaphoric Relations*.  
PhD thesis, MIT, 1981.
- [Aoun 85] Joseph Aoun.  
*A Grammar of Anaphora*.  
MIT Press. Linguistic Inquiry Monograph Eleven, 1985.
- [Bello y Cuervo 47] Bello and Cuervo, editors.  
*Esbozo de una Nueva Grammatica de la Lengua Espanola*.  
Real Academia Española, Comisión de Gramatica, 1947.
- [Borer 81] Hagit Borer.  
*Parametric Variation in Clitic Constructions*.  
PhD thesis, MIT, 1981.
- [Bouchard 83] Bouchard, D.  
*On The Content of Empty Categories*.  
Foris Publications, Dordrecht, Holland, 1983.
- [Boyer and Lapalme 84] Michel Boyer and Guy Lapalme.  
Generating Sentences from Semantic Networks.  
In *First International Workshop On Natural Language Understanding  
and Logic Programming*. IRISA, Université de Rennes, Campus de  
Beaulieu, 1984.

## [Boyer and Lapalme 85]

Michel Boyer and Guy Lapalme.  
 Generating Sentences from Semantic Networks.  
*Natural Language Understanding and Logic Programming*, Verónica Dahl and Patrick Saint-Dizier, editors.  
 North Holland, 1985.  
 Also found in the proceedings of the First International Workshop on Natural Language Understanding and Logic Programming, 1984, IRISA, Université de Rennes, Campus de Beaulieu, France.

## [Brown 86]

Charles Brown.  
*Using Gapping Grammars to Generate Spanish Clitics: a research proposal*.  
 Technical Report manuscript, School of Computing Science, Simon Fraser University, 1986.

## [Brown and Sempere 85]

Charles Brown and Juan Antonio Sempere-Martinez  
 On Clitics as Inflections.  
 In *Proceedings of the Second Annual Conference of the Northwest Linguistics Club, at the University of British Columbia*.  
 manuscript, 1985.

## [Brown and Sempere 87]

Charles Brown y Juan Antonio Sempere-Martinez.  
 Los clíticos del español vistos como inflexiones.  
*Revista Argentina de Lingüística* 3, 1987.

## [Brown, Dahl et al 86]

Charles Brown, Verónica Dahl, Diane Massam, Pierre Massicotte, T. Pattabhiraman.  
*Tailoring Government and Binding Theory for Automatic Natural Language Generation*.  
 Technical Report LCCR TR 86-4, Laboratory for Computer and Communications Research, School of Computing Science, Simon Fraser University, 1986.

## [Brown, Pattabhiraman et al 86]

Charles Brown, T. Pattabhiraman, Michel Boyer, Diane Massam, and Verónica Dahl.  
*Tailoring Conceptual Graphs for use in NL Translation*.  
 Technical Report LCCR TR 86-14, Laboratory for Computer and Communications Research, School of Computing Science, Simon Fraser University, 1986.

- [Brown, Pattabhiraman et al 87]  
 Charles Brown, T. Pattabhiraman, Michel Boyer, Diane Massam, and Verónica Dahl.  
 Tailoring Conceptual Graphs for use in NL Translation.  
 In *Conceptual Graphs in Knowledge Based Systems, Proceedings of the 1986 Thornwood Conference on Conceptual Graphs*, Norman Foo and John Sowa, editors. Addison Wesley, to appear August, 1987.
- [Chomsky 77] Noam Chomsky.  
*Essays on Form and Interpretation*.  
 North Holland, 1977.
- [Chomsky 82a] Noam Chomsky.  
*Lectures on Government and Binding, the Pisa Lectures, 2nd (revised) Edition*.  
 Foris Publications, Holland, 1982.
- [Chomsky 82b] Noam Chomsky.  
*Some Concepts and Consequences of the Theory of Government and Binding*.  
 Foris Publications, Holland, 1982.
- [Chomsky 86] Noam Chomsky.  
*Barriers, Linguistic Inquiry Monograph #13*.  
 MIT Press, 1986.
- [Coelho 79] H. Coelho.  
*A program conversing in Portuguese providing a library service*.  
 PhD thesis, University of Edinburgh, 1979.
- [Colmerauer 78] Alain Colmerauer.  
 Metamorphosis Grammars.  
*Lecture Notes in Computer Science, L. Bolc, editor*.  
 Springer-Verlag, 1978, pages 133-189.
- [Dahl 77] Verónica Dahl.  
*Un système déductif d'interrogation de banque de données en Espagnol*.  
 PhD thesis, Université d'Aix Marseille, 1977.  
 Thèse de Doctorat de Spécialité.
- [Dahl 79] Verónica Dahl.  
 Quantification in a Three Valued Logic for Natural Language Question-answering Systems.  
 In *Proceedings of 6th IJCAI, Tokyo*, 1979.

- [Dahl 80] Verónica Dahl.  
A Three Valued Logic for Natural Language computer applications.  
In *Proceedings of 10th International Symposium on Multiple Valued Logic, Illinois* . . . 1980.
- [Dahl 84] Verónica Dahl.  
More on Gapping Grammars.  
In *Proceedings Fifth Generation Computer Systems, Tokyo* . . . 1984.
- [Dahl 85] Verónica Dahl.  
*Logic Based Metagrammars for Natural Language Analysis*.  
Technical Report TR85-1, Simon Fraser University, 1985.
- [Dahl and Abramson 84] Verónica Dahl and Harvey Abramson.  
On Gapping Grammars.  
In *Proceedings Second International Logic Programming Conference, Uppsala, Sweden*, pages 77-88 Universitet Uppsala, 1984.
- [Dahl and Brown 86] Verónica Dahl and Charles Brown.  
*Constrained Discontinuous Grammars and Logic Programming*.  
Technical Report LCCR TR 86-17, Laboratory for Computer and Communications Research, School of Computing Science, Simon Fraser University, 1986
- [Dahl and McCord 83] Verónica Dahl and Michael McCord.  
Treating Coordination in Logic Grammars.  
*American Journal of Computational Linguistics* 9(2):69-91, 1983.
- [Dahl and Sambuc 76] Verónica Dahl and R. Sambuc.  
*Un système de bases de données en logique de premier ordre, en vue de sa consultation en langue naturelle*.  
Technical Report, DEA en Intelligence Artificielle, UER Marseille-Luminy, 1976.
- [Dahl and St-Dizier 86] Verónica Dahl and Patrick Saint-Dizier.  
*Constrained Discontinuous Grammars*.  
Technical Report LCCR TR 86-8, Laboratory for Computer and Communications Research, School of Computing Science, Simon Fraser University, 1986.



- [Dahl, Brown et al 86]  
 Veronica Dahl, Charles Brown, Michel Boyer, T. Pattabhiraman, and Diane Massam.  
*Mechanizing Expertise in Government and Binding Theory.*  
 Technical Report LCCR TR 86-10, Laboratory for Computer and  
 Communications Research, School of Computing Science, Simon Fraser  
 University, 1986.
- [Danlos 85] Laurence Danlos.  
 Un Survol des Recherches en Génération Automatique.  
*Revue Québécoise de Linguistique: Théorie Syntaxique* 15, 1985.
- [DeArmond 87] Richard C. DeArmond.  
*On the syntactic and lexical conceptual structure of Give.*  
 Technical Report, Dept. of Languages, Literature, and Linguistics, Simon  
 Fraser University, 1987.
- [F. Pereira 81] F. Pereira.  
 Extraposition Grammars.  
*American Journal of Computational Linguistics* 7:243-256, 1981.
- [F. Pereira and Warren 80]  
 Fernando C.N. Pereira and David H.D. Warren.  
 Definite clause grammars for language analysis.  
*Artificial Intelligence* 13(4):243-255, 1980.
- [Fillmore 68] C. J. Fillmore.  
 The Case for Case.  
*Universals in Linguistic Theory, E.Bach and R.Harms, editors.*  
 Holt, Rinehart, and Winston, NY, 1968.
- [Foo and Sowa 87]  
 Norman Foo and John Sowa, editors.  
*Conceptual Graphs in Knowledge Based Systems.*  
 Addison Wesley, (to appear) 1987.
- [Giannesini et al 85]  
 F. Giannesini, H. Kanoui, R. Pasero, et M. van Caneghem.  
*Prolog.*  
 InterEditions, Paris, 1985.
- [Gruber 65] J. S. Gruber.  
*Studies in Lexical Relations.*  
 PhD thesis, MIT, 1965.  
 Also published in [Gruber 76].

- [Gruber 76] J. S. Gruber.  
*Lexical Structures in Syntax and Semantics.*  
North Holland, 1976.
- [Hirschman 86] Lynette Hirschman.  
Conjunction in Met-Restriction Grammar.  
*Journal of Logic Programming* 4, 1986.
- [Hurtado 86] Hurtado, Alfredo.  
Clitic Chains, manuscript, published posthumously in a special issue  
*Homenaje A. Hurtado.*  
*Revista Argentina de Lingüística* 2, 1986.
- [Jackendoff 72] Ray S. Jackendoff.  
*Semantic Interpretation in Generative Grammar.*  
MIT Press, 1972.
- [Jackendoff 77] Ray S. Jackendoff.  
Constraints on Phrase Structure Rules.  
*Formal Syntax, P. W. Culicover et al, editors.*  
Academic Press, 1977. pages 249-283.
- [Jackendoff 83] Ray S. Jackendoff.  
*Semantics and Cognition.*  
MIT Press, 1983.
- [Jaeggli 80] Osvaldo Jaeggli.  
*On Some Phonetically Null Elements in Syntax.*  
PhD thesis, MIT, 1980.
- [Jaeggli 82] Osvaldo Jaeggli.  
*Topics in Romance Syntax.*  
Foris Publications, Dordrecht, 1982.
- [Joshi 87] Arvind Joshi.  
Word Order Variation in Natural Language Generation.  
In *Proceedings of the Sixth National Conference on AI*, pages 550-555.  
American Association for Artificial Intelligence, 1987.
- [Kaplan 79] Jerrold Kaplan.  
*Cooperative Responses from a Portable Natural Language DataBase  
Query System.*  
PhD thesis, Computer Science Department, University of Pennsylvania,  
1979.

- [Kayne 75] Richard S. Kayne.  
*French Syntax*.  
MIT Press, 1975.
- [Knowles 74] John Knowles.  
On acceptable Agrammaticality.  
*Linguistic Inquiry* 5, 1974.
- [Marcus 80] M. Marcus.  
*A Theory of Syntactic Recognition for Natural Language*.  
MIT Press, 1980.
- [May 77] May, Robert.  
*The Grammar of Quantification*.  
PhD thesis, MIT, unpublished, available through Indiana University  
Linguistics Club, 1977.
- [May 84] Robert May.  
*Logical Form: Its Structure and Derivation*.  
MIT Press, Cambridge, Massachusetts, 1984.
- [McCord 81] Michael McCord.  
Focalizers, the scoping problem, and semantic interpretation rules in logic  
grammars.  
*Logic Programming and its Applications*, D.H. Warren and M. van  
Canegham editors.  
Ablex Publishing, 1981.
- [McCord 82] Michael McCord.  
Using slots and modifiers in logic grammars for natural language.  
*Artificial Intelligence* 18:327-367, 1982.
- [McCord 84] Michael McCord.  
*A Syntax/Semantics System in PROLOG for the EPISTLE Project*.  
Technical Report, IBM T.J.Watson Research Center, 1984.
- [McCord 85] Michael McCord.  
Modular Logic Grammars.  
In *23 ACL*. 1985.
- [McKeown 83] K. McKeown.  
Focus constraints on language generation.  
In *Proceedings of the 8th International Joint Conference on Artificial  
Intelligence, Karlsruhe*. 1983.

- [Meisel 83] Jürgen. M. Meisel .  
Transfer as a Second-Language Strategy.  
*Language and Communication* 3(1), 1983.
- [Perlmutter 71] David M. Perlmutter.  
*Deep and Surface Constraints in Syntax*.  
Holt, N.Y., 1971.
- [Perlmutter 87] David M. Perlmutter.  
*Some Consequences of the Unaccusative Hypothesis for the Theory  
of Causatives*.  
Technical Report Seminar given at Simon Fraser University, 17 February,  
University of California, San Diego, 1987.
- [Perlmutter 83] David M. Perlmutter, editor.  
*Studies in Relational Grammar 1*.  
University of Chicago Press, 1983.
- [Perlmutter and Postal 83]  
David M. Perlmutter and H. Postal .  
Toward a Universal Characterization of Passivization.  
*Studies in Relational Grammar 1*.  
University of Chicago Press, 1983.
- [Perlmutter and Rosen 84]  
David M. Perlmutter and Carol Rosen, editors.  
*Studies in Relational Grammar 2*.  
University of Chicago Press, 1984.
- [Pfaff 79] C. Pfaff.  
Constraints on Language Mixing: intrasentential code-switching and  
borrowing in Spanish/English.  
*Language* 55, 1979.
- [Pique 82] Jean Francois Pique.  
On a semantic representation of natural language sentences.  
In *Proceedings of the First International Logic Programming  
Conference*. Faculté des Sciences de Luminy, Marseille, 1982.
- [Poplack 80] Shana Poplack.  
Sometimes I'll Start a Sentence in English y termino en español.  
*Linguistics* 18, 1980.
- [Popowich 85] Fred Popowich.  
Unrestricted Gapping Grammars.  
In *Proceedings 9th International Joint Conference on Artificial  
Intelligence, Los Angeles*. AAAI, 1985.

- [Porto and Filgueiras 84] António Porto and Miguel Filgueiras.  
Natural Language Semantics: a logic programming approach.  
*IEEE Computer* 3, 1984.
- [Quicoli 76] A. Carlos Quicoli.  
Conditions on Clitic Movement in Portuguese.  
*Linguistic Analysis* 6(2), 1976.
- [Radford 81] Andrew Radford.  
*Transformational Syntax*.  
Cambridge University Press, 1981.
- [Rivas 77] Alberto Rivas.  
*A Theory of Clitics*.  
PhD thesis, MIT, 1977.
- [Roberge 85] Yves Roberge.  
*Subject Doubling, Free Inversion, and Null Argument Languages (PhD Thesis. to appear)*.  
Technical Report. University of British Columbia, 1985.
- [Roldan 73] Mercedes Roldán.  
Constraints on Clitic Insertion in Spanish.  
*Linguistic Studies in Romance Languages. Proceedings of the Third Linguistic Symposium in Romance Languages*.  
Georgetown University Press, Washington, DC, 1973.
- [Sabatier 84] Paul Sabatier.  
Puzzle Grammars.  
In *Proceedings Conference on Logic Programming and Natural Language Processing, Rennes*. 1984.
- [Safir 85] K. Safir.  
*Syntactic Chains*.  
Cambridge University Press, 1985.
- [Saint-Dizier 84] Patrick Saint-Dizier.  
Handling Quantifier Scoping Ambiguities in a Semantic Representation of Natural Language Sentences.  
*Natural Language Understanding and Logic Programming*. Verónica Dahl and Patrick Saint-Dizier, editors.  
North-Holland, 1984.

- [Saint-Dizier 86a] Patrick Saint-Dizier.  
An Approach to NL Semantics in Logic Programming.  
*Journal of Logic Programming* 4, 1986.
- [Saint-Dizier 86b] Patrick Saint-Dizier.  
*Context Dependent Determiners in Logic Programming: Semantic Representations and Priorities.*  
Technical Report, IRISA, Campus de Beaulieu, Rennes, France, 1986.
- [Saint-Dizier 87] Patrick Saint-Dizier.  
*Contextual Discontinuous Grammars.*  
Technical Report 332, to appear in the proceeding of the Second International Workshop on Natural Language Understanding and Logic Programming, Simon Fraser University, INRIA, 1987.
- [Selkirk 82] Elizabeth O. Selkirk.  
*The Syntax of Words.*  
MIT Press, 1982.
- [Sharp 85] Randy Sharp.  
*A Government-Binding Approach to NL Translation.*  
Technical Report, MSc thesis, UBC, Dept. of Computer Science, 1985.
- [Simmons and Slocum 72] R. F. Simmons and J. Slocum.  
Generating English Discourse from Semantic Networks.  
*CACM* 15(10), 1972.
- [Sowa 84] John Sowa.  
*Conceptual Structures: Information processing in mind and machine.*  
Addison-Wesley, 1984.
- [Stabler 87] Edward P. Stabler.  
Restricting Logic Grammars with Government Binding Theory.  
*Computational Linguistics* 13(1-2):1-10, 1987.
- [Stowell 81] Timothy Stowell.  
*Origins of Phrase Structure.*  
PhD thesis, MIT, 1981.
- [Strozer 76] Judith Reina Strozer.  
*Clitics in Spanish.*  
PhD thesis, UCLA, 1976.

- [Westfal 79] Germán Westfal.  
*Subjects and Pseudo-Subjects in Spanish: The verb agreement question in the impersonal se construction.*  
Linguistic Research Inc., Edmonton, 79.